Models of response inhibition in the stop-signal and stop-change paradigms

Frederick Verbruggen a,b,*, Gordon D. Logan a,1

1 Department of Psychology, Vanderbilt University, Nashville, TN 37203, United States
2 Department of Experimental Psychology, Ghent University, Belgium

ABSTRACT

The stop-signal paradigm is very useful for the study of response inhibition. Stop-signal performance is typically described as a race between a go process, triggered by a go stimulus, and a stop process, triggered by the stop signal. Response inhibition depends on the relative finishing time of these two processes. Numerous studies have shown that the independent horse-race model of Logan and Cowan [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, 295–327] accounts for the data very well. In the present article, we review the independent horse-race model and related models, such as the interactive horse-race model [Boucher, L., Palmeri, T.J., Logan, G.D., Schall, J.D., 2007. Inhibitory control in mind and brain: an interactive race model of countermanding saccades. Psychological Review 114, 376–397]. We present evidence that favors the independent horse-race model but also some evidence that challenges the model. We end with a discussion of recent models that elaborate the role of a stop process in inhibiting a response.

© 2008 Elsevier Ltd. All rights reserved.
Response inhibition is considered to be a key component of executive control (e.g., Andres, 2003; Aron, 2007; Logan, 1985a; Miyake et al., 2000; Stuphorn and Schall, 2006). The concept refers to the ability to suppress responses that are no longer required or inappropriate, which supports flexible and goal-directed behavior in ever-changing environments. In everyday life, there are many examples of the importance of response inhibition, such as stopping yourself from crossing a street when a car comes around the corner without noticing you. Response-inhibition deficits have also been linked to disorders such as attention-deficit/hyperactivity disorder (e.g., Nigg, 2001; Oosterlaan et al., 1998; Schachar and Logan, 1990), obsessive-compulsive disorder (e.g., Chamberlain et al., 2006; Menzies et al., 2007; Penades et al., 2007), and substance abuse disorders (e.g., Monterosso et al., 2005; Nigg et al., 2006). Response-inhibition deficits are discussed in more detail by Chambers et al. (submitted for publication) and by Jentsch et al. (submitted for publication) in this issue. A paradigm that is most suitable for the investigation of response inhibition in a laboratory setting is the stop-signal paradigm (Lappin and Eriksen, 1966; Logan and Cowan, 1984; Vince, 1948).

In the standard stop-signal paradigm, subjects perform a choice reaction time (RT) task (i.e., the go task; also referred to as the primary task), such as responding to the shape of a stimulus (e.g., press a left key for a square and a right key for a circle). Occasionally, the go stimulus is followed by an auditory tone (i.e., the stop signal), which instructs subjects to withhold their response. Fig. 1 depicts an example of the trial course of a typical stop-signal experiment. Typically, subjects can inhibit their response when the stop signal is presented close to the moment of stimulus presentation, but they cannot inhibit their response when the stop signal is presented close to the moment of response execution. To account for these observations, Logan (1981) and Logan and Cowan (1984) proposed a race between a go process and a stop process and argued that the relative finishing time of these two processes determines whether subjects will respond or stop. In this article, we will present a theoretical review of the so-called independent horse-race model and related models, and we will discuss the most important measures of inhibitory control in the stop-signal paradigm.

1. Early horse-race models of response inhibition

The idea that response inhibition depends on the relative finishing time of a go process and a stop process has always dominated the stop-signal literature. Vince (1948) showed that subjects could stop their response only when the delay between the go stimulus and the stop signal (stop-signal delay or SSD) was short (i.e., 50 ms). At longer delays (i.e., 100 ms and longer), response inhibition was very rare, which suggests that the stop process started too late to cancel the response. Lappin and Eriksen (1966) also manipulated stop-signal delay. They argued that when the stop signal is delayed, subjects delay the go process in order to keep the probability of responding [p(respond|signal)] similar across delays.

The race idea was present implicitly in the work of Vince (1948), and Lappin and Eriksen (1966). Ollman (1973) formalized the idea of a race between a go process and a stop process. He used the stop-signal paradigm to test the hypothesis that subjects perform choice reaction tasks by setting a subjective deadline and then making either a stimulus-controlled response or a “guess” response, depending on whether stimulus-controlled processing finished before the deadline. In the stop-signal task, subjects would set the deadline so that the stop signal could be detected before the deadline (i.e., \(T_{d\text{-go}} + D > T_{d\text{-stop}} + SSD\)), where \(T_{d\text{-go}}\) is the time needed to detect the go stimulus, \(D\) is the subjective deadline, and \(T_{d\text{-stop}}\) is the time needed to detect the stop signal. When the stop signal is detected before the deadline (i.e., \(T_{d\text{-go}} + D > T_{d\text{-stop}} + SSD\)), subjects successfully stop the response; when the stop signal is detected after the deadline (i.e., \(T_{d\text{-go}} + D < T_{d\text{-stop}} + SSD\)), subjects erroneously emit the response. Ollman implemented this deadline model with specific assumptions about the parametric form of the finishing-time distributions (i.e., he assumed normal and exponential distributions), but rejected it based on the behavioral data. Nevertheless, the idea that stopping a response depends on the relative finishing time of the go process and the stop process was established.

2. The independent horse-race model

Early horse-race models mainly focused on describing go and stop performance either qualitatively (Lappin and Eriksen, 1966) or with narrowly focused quantitative assumptions (Ollman, 1973). These models were limited in generality and lacked a precise description of the main variables of interest, namely the difficulty of the stop process and the latency of the stop process (stop-signal reaction time or SSRT). Unlike the latency of an overt choice response, the latency of the response to the stop signal (i.e., suppressing the go response) cannot be measured directly. Later versions of the horse-race model dealt with this issue. Consistent with earlier ideas of Lappin and Eriksen (1966) and Ollman (1973), Logan (1981) suggested that performance in the stop-signal paradigm can be modeled as a “horse race” between a go process, which is triggered by the presentation of a go stimulus, and a stop process, which is triggered by the presentation of the stop signal. When the stop process finishes before the go process (i.e., \(go > SSRT + SSD\)), response inhibition is successful and no response is emitted (signal-inhibit; see Fig. 2A); when the go processes finishes before the stop process (i.e., \(go < SSRT + SSD\)), response inhibition is unsuccessful and the response is incorrectly emitted (signal-respond; see Fig. 2A). Logan (1981) used this horse-race idea to estimate the
Fig. 2. (A) Graphic representation of the horse-race idea. The length of the bars represents the duration of the process (SSD = stop-signal delay; SSRT = stop-signal reaction time). (B) Graphic representation of the assumptions of the independent horse-race model of Logan and Cowan (1984), indicating how the probability of responding \( p(\text{respond|signal}) \) and the probability of inhibiting \( p(\text{inhibit|signal}) \) depend on the distribution of go reaction times, stop-signal delay (SSD) and stop-signal reaction time (SSRT).

'Unobservable' latency of the stop process (i.e., SSRT) and this has contributed to the increasing popularity of the paradigm.

Logan and Cowan (1984) developed a general formal version of the horse-race model that described both go and stop performance in terms of relations between the finishing-time distributions of stop and go processes. Their derivations did not depend on the specific parameterization of the finishing-time distributions (cf. Ollman, 1973). They made parameter-free predictions and developed parameter-free measures that would hold regardless of the functional form of the finishing-time distributions. They made certain assumptions about the independence between the go and stop process (see below), which allowed them to successfully account for inhibition functions, stop-signal reaction times and RTs for trials that escaped inhibition (i.e., signal-respond RTs) in a wide variety of data sets.

2.1. Inhibition functions

Inhibition functions depict the relation between \( p(\text{respond|signal}) \) and SSD. These functions are important theoretically because they reflect the outcome of the race between the go process and the stop process (Logan and Cowan, 1984). They are important empirically because they reflect the ability to control responses; they can be used to compare inhibitory control in different groups, tasks or conditions. However, we will show that differences in inhibition functions can be due to several factors, making it difficult to interpret them sometimes.

2.1.1. Inhibition functions: the basics

Inhibition functions are influenced primarily by three factors: SSD, the go RT distribution and the SSRT distribution. Logan and Cowan (1984) described the relations between these factors formally but noted that it is easier to see the relations if SSRT is treated as a constant instead of a random variable. Mathematical analyses (Logan and Cowan, 1984) and Monte Carlo simulations (Band et al., 2003; De Jong et al., 1990) showed that assuming SSRT is constant does not systematically bias the SSRT estimates for most estimation methods (see below). Therefore, we adopt this assumption below for ease of exposition.

The relation between \( p(\text{respond|signal}) \), SSD, go RT, and SSRT is depicted in Fig. 3. The independent horse-race model assumes that the SSD will influence the relative finishing time of the stop process: when SSD increases, the stop process will start later and therefore, finish later compared to the go process. Consequently, the probability that the go process will finish before the stop process increases and response inhibition will succeed less often. This relation between SSD and \( p(\text{respond|signal}) \) is depicted in Fig. 3A. \( p(\text{respond|signal}) \) is represented by the area under the curve to the left of each dashed line in the left panel of Fig. 3A. As can be seen, the response to the stop signal cuts off more of the go RT distribution when SSD increases. In practice, \( p(\text{respond|signal}) \) will be 0 when the stop signal occurs early enough (i.e., SSD + SSRT < the shortest go RT of the empirical distribution). \( p(\text{respond|signal}) \) will be 1 when the stop signal occurs late enough (i.e., SSD + SSRT > the longest go RT of the empirical distribution).

In theory (and in practice), \( p(\text{respond|signal}) \) increases monotonically from 0 to 1 as SSD increases, tracing out the inhibition function (see the right panel of Fig. 3A).

Differences in go RT and SSRT will also influence \( p(\text{respond|signal}) \). For every SSD, \( p(\text{respond|signal}) \) will decrease when the go RT increases (i.e., when the distribution is shifted to the right) because the probability that the stop process finishes before the go process increases (see Fig. 3B). Thus, subjects can keep \( p(\text{respond|signal}) \) similar across delays by slowing go RT appropriately (e.g., Lappin and Eriksen, 1966; Logan, 1981). This is shown in the right panel of Fig. 3B. For every SSD, \( p(\text{respond|signal}) \) will increase when SSRT increases because the probability that the stop process will finish after the go process increases. This is shown in Fig. 3C. Aligning inhibition functions

The independent horse-race model predicts that \( p(\text{respond|signal}) \) depends on the relative finishing time of the go process and stop process and not on their relative starting times. \( p(\text{respond|signal}) \) will be the same for different conditions even though SSD, SSRT and the underlying go RT distribution may be different, provided that the relative finishing time of the go process and the stop process is the same (see Fig. 3A–C; Logan and Cowan, 1984). Therefore, empirical inhibition functions for different subjects, tasks, or conditions may be aligned by plotting \( p(\text{respond|signal}) \) against relative finishing time, but they may be misaligned when \( p(\text{respond|signal}) \) is plotted against SSD because SSD reflects the relative starting time of the go process and the stop process. When inhibition functions are plotted against relative starting time (i.e., SSD), they are shifted to the right when go RT increases (see Fig. 3B) and shifted to the left when SSRT increases (see Fig. 3C).

The independent horse-race model predicts that go RT differences can be taken into account by plotting \( p(\text{respond|signal}) \) against \( RT_{\text{go}} - \text{SSD} \). For example, the two inhibition functions in the right panel of Fig. 3B would be aligned if we plotted \( p(\text{respond|signal}) \) against \( RT_{\text{go}} - \text{SSD} \). A second alignment method also considers variability in go RT. As shown in Fig. 4, an inhibition function can be influenced by variability in go RT, even when the mean go RT remains the same (for ease of exposition, we depicted two normal go RT distributions in Fig. 4, but the same principles apply to...
skewed distributions that are more commonly observed in the literature). When variability increases, a smaller proportion of the go RT distribution will fall between two consecutive SSDs (see the left panel of Fig. 4). Consequently, the inhibition function for the condition with the greater variability would be flatter than the condition with the lesser variability (see the right panel of Fig. 4). Logan et al. (1984) took RT variance (SD\(_{\text{go}}\)) into account by plotting \(p(\text{respond}|\text{signal})\) against \((\text{RT}_{\text{go}} - \text{SSD})/\text{SD}_{\text{go}}\). However, they found that this second method did not improve the fit substantially compared to the first method that did not take SD\(_{\text{go}}\) into account.

Misalignment could also be due to differences in SSRT (see Fig. 3C). Logan et al. (1984) proposed a third alignment method that takes differences in mean go RT, \(\text{SD}_{\text{go}}\), and SSRT into account and plots inhibition functions in terms of a Z score (see e.g., Armstrong and Munoz, 2003; Logan and Cowan, 1984; Logan et al., 1984; Logan and Irwin, 2000; Schachar and Logan, 1990; van der Schoot et al., 2000). This method plots \(p(\text{respond}|\text{signal})\) against the relative finishing time (RFT) of the go and stop process in standard deviation units, resulting in a Z score: \(\text{ZRFT} = (\text{RT}_{\text{go}} - \text{SSD} - \text{SSRT})/\text{SD}_{\text{go}}\). Alignment is typically best for the ZRFT functions, although the difference with the first alignment method (i.e., \(\text{RT}_{\text{go}} - \text{SSD}\)) is not

\[\text{ZRFT} = (\text{RT}_{\text{go}} - \text{SSD} - \text{SSRT})/\text{SD}_{\text{go}}.\]
always large, suggesting that differences in mean go RT are more important than differences in go RT variance and SSRT (Logan and Cowan, 1984). However, these tests of alignment are usually based on visual inspection. Currently, there are no quantitative methods available for evaluating alignment.

If the inhibition functions can be aligned after plotting them as a function of RTgo – SSD or ZRFT, then one can conclude that the same inhibitory processes (i.e., the independent horse-race model) apply to them (Logan and Cowan, 1984). For example, Logan and Irwin (2000) showed that the same inhibitory principles apply to the inhibition of eye and hand movements, even though go RT and SSRT were shorter for eye movements than for hand movements (see below). However, it may not always be possible to align different inhibitory functions. Failures of alignment suggest that the independent horse-race model does not apply to one or more of the inhibition functions (Logan and Cowan, 1984). Misalignment of ZRFT functions could indicate that the inhibition mechanism is triggered less often or that it is substantially more variable (Schachar and Logan, 1990; Tannock et al., 1995). Misalignment of ZRFT functions could also indicate the presence of ballistic components in processing that cannot be controlled (see below). However, it is difficult to distinguish between these possibilities (see Logan, 1994). Moreover, Band et al. (2003) showed that variability in go RTs can still influence the shape of ZRFT functions, even though these functions are intended to account for variability in go RTs.

### 2.1.3. Interim conclusions

Inhibition functions depict the outcome of the race between the go process and the stop process. Inhibition functions are influenced by SSD, go RT, SDgo, and SSRT, and several methods have been proposed to align inhibition functions in different tasks, populations or conditions. Misalignment of ZRFT functions can suggest differences in the inhibitory processes or the presence of ballistic components in the go process. However, Monte Carlo simulations of Band et al. (2003) suggest that differences in inhibition functions should be interpreted carefully because it is not always entirely clear what factors are causing the misalignment.

### 2.2. Stop-signal reaction times

Most stop-signal studies focus mainly on the latency of the stop process – SSRT – as an index of inhibitory control. Unlike go RT, SSRT cannot be measured directly but it can be estimated from methods that are based on the assumptions of the independent horse-race model (Logan and Cowan, 1984). Several methods for estimating SSRT are available in the literature (e.g., Colonius, 1990; De Jong et al., 1990; Logan, 1981; Logan and Cowan, 1984; for a review, see Logan, 1994). The strongest methods assume SSRT is a random variable. However, to simplify the presentation, we assume that SSRT is constant in the following paragraph.

SSRT corresponds to the time interval between the point at which the stop process starts (i.e., when the stop signal is presented) and the point at which the stop process finishes (see Fig. 2B). The independence horse-race model assumes that the stop process starts when the stop signal is presented (which is known to the experimenter, based on the SSD). The point at which the stop process finishes can be estimated on the basis of the observed go RT distribution on no-signal trials and the observed p(respond|signal) for a given SSD. More specifically, the point at which the internal response to the stop signal occurs (RTir) is estimated by integrating the go RT distribution and finding the point at which the integral equals p(respond|signal) (Logan, 1981, 1994; Logan and Cowan, 1984). In Fig. 2B, this corresponds to finding the point at which the dashed line crosses the time axis. Once the starting point and the finishing point are known, SSRT can be estimated by subtracting SSD from RTir. In practice, there are several methods to estimate SSRT (for reviews, see Band et al., 2003; Logan, 1994). Which estimation method is most suitable depends on how SSD is set. There are two basic procedures for setting SSD: (1) using a variable number of fixed SSDs (i.e., the fixed-SSDs procedure) or (2) adjusting SSD dynamically (i.e., the tracking procedure).

#### 2.2.1. SSRT estimation for the fixed-SSDs procedure

When fixed delays are used, SSRT can be estimated in several ways. The most commonly used method is probably the integration method (Logan and Cowan, 1984), which assumes that SSRT is a constant. The integration method was described in principle in the preceding paragraph: SSRT is estimated by subtracting SSD from the finishing time of the stop process. The finishing time is determined by integrating the go RT distribution. In practice, the no-signal go RTs are rank-ordered, then the nth RT is selected, where n is obtained by multiplying the number of no-signal go RTs in the distribution by the probability of responding at a given delay. For example, when there are 100 no-stop-signal trials, and p(respond|signal) = .25 at a given delay, then nth RT is the 25th fastest go RT. To estimate SSRT, SSD is subtracted from the nth RT. This process is repeated for each SSD for each subject. The results are typically averaged across SSDs. The advantage of this method is that it allows SSRT estimates for every SSD. However, we will show (see Section 3.1.3) that the integration method may be more susceptible to violations of the assumptions of the independent horse-race model than other estimation methods.

A second method for estimating SSRT is the mean method. The mean method assumes that SSRT is a random variable. SSRT is estimated by determining the mean of the inhibition function, which is then subtracted from the mean go RT (see e.g., Logan and Cowan, 1984). In practice, the mean of the inhibition function can be estimated from the observed probabilities of responding at each SSD. If there is a complete inhibition function [i.e., ranging from p(respond|signal) = 0 to p(respond|signal) = 1], then the mean of the inhibition function is the mean of the inhibition function (e.g., Chambers et al., 2006; Hanes and Schall, 1995), which amounts to making parametric assumptions about the form of the inhibition function. The median of the inhibition function is determined, SSRT is estimated by subtracting the mean of the inhibition function from the mean go RT. Others have used the median of the inhibition function (i.e., the median method) to estimate SSRT by fitting a logistic or Weibull function to the inhibition function (e.g., Chambers et al., 2006; Hanes and Schall, 1995). If the distribution is symmetrical, the median and the mean are the same, so using the mean or median of the inhibition function yields the same result (Logan and Cowan, 1984).

When multiple SSDs are used, the distribution of SSRTs can be estimated from the go RT distribution of no-signal trials and the go RT distribution of signal–respond trials (e.g., Colonius, 1990; De Jong et al., 1990). However, this estimation method requires a larger number of observations than the integration, mean or median methods because it is very sensitive to the quality of the data, particularly at the tails of the distributions (Logan, 1994). Moreover, it seems to underestimate SSRT, compared to the other estimation methods (Band et al., 2003).
2.2.2. SSRT estimation for the tracking procedure

SSD can also be set dynamically. Several dynamic methods have appeared in the literature (e.g., Logan et al., 1984; Schachar et al., 1995). A common method involves adjusting SSD after every trial (i.e., the one-up one-down procedure; see e.g. Logan et al., 1997; Osman et al., 1986). After successful stopping, SSD increases, which handicaps the stop process on the next stop-signal trial. After unsuccessful stopping, SSD decreases, which handicaps the go process on the next stop-signal trial. If the increases and decreases in SSD on each trial are equal in magnitude, the tracking procedure should result in an overall p(respond|signal) of .50 (e.g., Aрон and Poldrack, 2006; Logan et al., 1997; Ridderinkhof et al., 1999; Scheres et al., 2001; Verbruggen et al., 2004). Thus, the tracking procedure compensates for differences between and within subjects, resulting in a similar p(respond|signal) for different subjects, tasks or conditions.

A major advantage is that SSRT can be estimated easily with the mean method when the tracking procedure produces p(respond|signal) equal to .50. Logan and Cowan (1984) showed that SSRT can be calculated by subtracting the mean of the inhibition function from the mean RT. When the tracking procedure produces p(respond|signal) equal to .50, the mean of the inhibition function is equal to the mean SSD. Consequently, SSRT can be estimated by subtracting the observed mean SSD from the observed mean go RT (e.g., Logan et al., 1997). Both simulations (Band et al., 2003) and reliability tests showed that when the tracking procedure is used, SSRTs estimated with the mean method are most reliable (Logan et al., 1997; Williams et al., 1999). However, researchers are advised to estimate SSRT differently when p(respond|signal) is significantly different from .50 (see Verbruggen et al., 2008b).

Logan et al. (1997) used the mean method in combination with the tracking procedure. Others have used the integration method in combination with the tracking procedure to estimate SSRT. Variants of the integration method involve estimating SSD by subtracting mean SSD from the median RT (e.g., Aрон and Poldrack, 2006; Cohen and Poldrack, 2008) or subtracting mean SSD from the nth RT (Ridderinkhof et al., 1999; Schachar et al., 2007; van den Wildenberg et al., 2002; Verbruggen et al., 2004), where the nth RT is determined by multiplying the number of RTs in the go RT distribution by the overall p(respond|signal) (note that the two methods will yield the same SSRT when overall p(respond|signal) equals .50). The Monte Carlo simulations of Band et al. (2003) showed that the integration method resulted in reliable SSRT estimates for central SSDs (i.e., SSDs for which p(respond|signal) is close to .50). However, there were no explicit tests of the reliability of the integration method in combination with the tracking procedure.

2 The p(respond|signal) can be manipulated by adjusting SSD differently after successful and unsuccessful inhibition. If SSD decreases after every signal-respond trial and increases after every two signal-inhibit trials (the one-down two-up procedure), the tracking procedure typically results in an overall p(respond|signal) of approximately .29; if SSD decreases after every two signal-respond trials and increases after every signal-inhibit trial (the two-down one-up procedure), the tracking procedure typically results in an overall p(respond|signal) of approximately .71 (Osman et al., 1986). These numbers correspond to what is predicted based on psychometric functions (for a review on adaptive procedure in psychological research, see e.g., Leek, 2001).

2.2.3. Interim conclusions

One of the most important contributions of the independent horse-race model of Logan and Cowan (1984) is that it provides theoretically justified estimates of the covert latency of the stop process. There are several methods to estimate SSRT, but simulations of Band et al. (2003) showed that SSRT estimates that are derived from the central part of the distribution (i.e., for SSDs for which p(respond|signal) approaches .50) are most reliable. These central estimates are less influenced by variability in go RT and SSRT and relatively robust against violations of the independence assumptions (see below). Thus, methods that use the mean method or the median method typically result in more reliable SSRT estimates than the integration method. The integration method will result in reliable SSRT estimates only for those SSDs for which p(respond|signal) approximates .50.

SSD can be set via the fixed-SSDs procedure or the tracking procedure. Both methods can result in reliable SSRT estimates. However, fewer stop-signal trials are needed for the tracking procedure than for the fixed-SSDs procedure (Band et al., 2003). Thus, the tracking procedure may be preferred over the fixed-SSDs procedure from the perspective of experimental economy. Based on simulations and reliability tests (Band et al., 2003; Williams et al., 1999), we advise researchers to have at least 40–50 stop signals when the tracking procedure is used. Given that stop-signals are typically presented on 25% of the trials, this implies that approximately 160–200 trials are needed to obtain reliable SSRT estimates.

2.3. Measures of inhibitory control in practice

Inhibition functions and SSRTs have been used to investigate response inhibition in cognitive psychology, lifespan development, psychopathology, and cognitive neuroscience (Verbruggen and Logan, 2008c). In this section, we will discuss a selective subset of stop-signal studies to demonstrate how inhibition functions and SSRTs can be used to study inhibitory control in practice.

In cognitive psychology, inhibition functions and SSRTs have been used to study inhibitory control in discrete and continuous tasks, such as simple and two-choice RT tasks (Logan et al., 1984), typewriting (Logan, 1982), simple pursuit tasks (e.g., Morein-Zamir and Meiran, 2003; Morein-Zamir et al., 2004) or tasks that required subjects to make arm movements (e.g., McGarry and Franks, 1997; Mirabella et al., 2006). In general, these studies showed that inhibition functions and SSRTs were comparable for different tasks and response types. Inhibitory control is also comparable for different effectors. Xue et al. (2008) showed that SSRTs were similar for interrupting speech and interrupting a manual key response. However, SSRT is typically shorter for eye movements than for hand movements (e.g., Boucher et al., 2007b; Logan and Irwin, 2000), although the aligned inhibition functions for hand and eye movements suggest that inhibitory control operates according to the same horse-race model principles (Logan and Irwin, 2000). All of these cognitive studies used inhibition functions and SSRT to determine whether the same inhibitory control mechanisms can be generalized to different tasks or different effectors. Other cognitive studies focused on the factors that influence inhibitory control in a single task. For example, several studies compared inhibition functions and SSRTs to examine how response inhibition is influenced by stop-signal modality and intensity (e.g., Assress and Carpenter, 2001; Cabel et al., 2000; Morein-Zamir and Kingstone, 2006; van der Schoot et al., 2005), response properties (e.g., van den Wildenberg et al., 2003), or by the presentation of various types of distracting information (e.g., Ridderinkhof et al., 1999; Verbruggen and De Houwer, 2007; Verbruggen et al., 2004). We will come back to the effect of distracting information later.

Inhibitions functions and SSRT estimates have also been used extensively to study response inhibition in many other literatures, such as lifespan development, clinical psychology and psychopathology (see also Chambers et al., submitted for publication; Jentsch et al., submitted for publication). For example, several
studies have demonstrated that SSRT is elevated in younger children (e.g., van den Wildenberg and van der Molen, 2004; Williams et al., 1999) and older adults (e.g., Kramer et al., 1994; Rush et al., 2006), compared to young adults. Moreover, a comparison of go RT and SSRT showed that go and stop performance develop and decline independently. Psychopathologists used SSRT to study inhibitory deficits clinical populations, such as impulsive people (e.g., Logan et al., 1997; Stahl and Gibbons, 2007) and patients with obsessive-compulsive disorder (e.g., Chamberlain et al., 2006; Penades et al., 2007). One of the most replicated findings is that SSRT is elevated in children with attention deficit/hyperactivity disorder (ADHD) (e.g., Jennings et al., 1997; Nigg, 1999; Schachar and Logan, 1990; Schachar et al., 2000) compared to control groups (for a review, see Lijffijt et al., 2005). Adults with ADHD also show longer SSRTs compared to control groups (e.g., Aron et al., 2003a; Ossmann and Mulligan, 2003). Several studies showed that inhibition functions for ADHD groups and control groups could not be aligned (e.g., Pliszka et al., 1997; Schachar and Logan, 1990; Tannock et al., 1995). According to these authors, the misaligned inhibition functions suggest that the inhibition mechanism of the ADHD group was triggered less frequently or was substantially more variable (see e.g., Schachar and Logan, 1990). However, the Monte Carlo simulations of Band et al. (2003) suggest that other factors, such as variability in go RT, could also have contributed to the misalignment (even though 2RF functions are supposed to take \( SD_{\text{r}} \) into account).

SSRT estimates have also been used to study inhibitory control in the brain (see also Chambers et al., submitted for publication). The most prominent example of how SSRT is used by cognitive neuroscientists comes from single-unit recording studies by Hanes, Schall and colleagues (e.g., Hanes et al., 1998; Ito et al., 2003; Pare and Hanes, 2003; Stuphorn et al., 2000) who use SSRT as a criterion for determining whether or not different types of neurons can contribute to the control of eye movements. Neurons whose activity modulates after a stop signal but before SSRT can contribute to movement control; neurons whose activity modulates after SSRT cannot contribute. By this criterion, movement-related but not visually responsive neurons in frontal eye fields (Hanes et al., 1998) and superior colliculus (Pare and Hanes, 2003) contribute to movement control, but neurons in supplementary eye fields (Stuphorn et al., 2000) and anterior cingulate cortex (Emerick et al., 2008; Ito et al., 2003) do not directly contribute to movement control. Instead, neurons in supplementary eye fields and anterior cingulated cortex would be involved in monitoring of go and stop performance.

Another prominent example of the use of SSRT in cognitive neuroscience research comes from FMRI and lesion studies. Several studies showed that frontal regions, such as right inferior frontal gyrus (IFG) and pre-supplementary motor area (pre-SMA) are involved in response inhibition. To elucidate the role of these regions, Aron and colleagues have related SSRT to the activation in these brain regions. They showed that SSRT was negatively correlated to right IFG activation (i.e., more activation was associated with shorter SSRTs; Aron et al., 2007; Aron and Poldrack, 2006) but not to pre-SMA activation (Aron et al., 2007). Lesion studies showed that SSRT is correlated to the degree of right IFG damage in patients (Aron et al., 2003b). van den Wildenberg and colleagues demonstrated that deep-brain stimulation of the subthalamalic nucleus improved both go and stop performance in patients with Parkinson, but a comparison of go RT and SSRT showed that the effects of STN on go and stop performance may be functionally independent (van den Wildenberg et al., 2006).

In sum, inhibition functions and SSRTs have proven useful in practice. These two measures have allowed researchers to study inhibitory control in a variety of populations, tasks and situations and shed further light on how a response can be inhibited in both healthy and clinical groups.

2.4. Ballistic stages of controlled processing and the point-of-no-return

In the previous sections, we showed how the independent horse-race model of Logan and Cowan (1984) captures the difficulty and latency of control in the stop-signal paradigm. In addition, the independent horse-race model addressed the measurement of the ballistic component of control. Ballistic processes are processes that must run to completion once they have been launched, and therefore, cannot be inhibited. By contrast, controlled processes can be inhibited at any point (e.g., Zbrodoff and Logan, 1986). The temporal boundary between controlled processing stages and ballistic processing stages is called the point-of-no-return.

The independent horse-race model addresses ballistic processing by assuming two stages in the go process: a controlled stage with duration \( RT_c \) and a ballistic stage with duration \( RT_b \) (Logan and Cowan, 1984; Osman et al., 1986). The controlled go race processes with the stop process. If the stop process finishes before the controlled go process (i.e., \( RT_c > SSRT + SSD \)), subjects inhibit their responses. If the controlled go process finishes before the stop process (i.e., \( RT_c < SSRT + SSD \)), subjects fail to inhibit their responses. Inhibition fails whenever \( RT_c < SSRT + SSD \), even if total RT is longer than SSRT + SSD (i.e., \( RT_c + RT_b > SSRT + SSD \)). Consequently, factors that affect controlled stage should influence go RT and the inhibition function by the same amount. By contrast, factors that affect ballistic stages should influence go RT but not the inhibition function because these factors influence response stages that are beyond cognitive control (Logan, 1981; Logan and Cowan, 1984; Osman et al., 1986). Several studies used inhibition functions and SSRTs to study controlled and ballistic stages in reaction time tasks (e.g., Cavina-Pratesi et al., 2004; Cohen and Poldrack, 2008; Logan, 1981; Osman et al., 1986, 1990). These studies showed that most experimental manipulations, such as stimulus discriminability and stimulus–response compatibility, affect stages before the point-of-no-return (but see Osman et al., 1986, Experiments 2 and 3). For example, the go RT distribution and the inhibition function shifted to the right by the same amount when the stimuli were more difficult to discriminate, which suggests that stimulus discriminability affected controlled stages (Logan, 1981). The findings suggested that the ballistic stages contribute only a small part to the overall go RT. This is consistent with mathematical analyses of Logan and Cowan (1984), which showed that the ballistic stages of the go process must be very brief. This idea was further supported by studies that measured electromyograms, which showed that subjects could still inhibit responses that produced electromyographic responses (De Jong et al., 1990). Combined, these findings suggest that ballistic processing stages must be very late and very brief in duration if they exist at all (e.g., De Jong et al., 1990; Gao and Zelaznik, 1991; McCrary and Franks, 1997; McCrary et al., 2000; Osman et al., 1990). These studies focused on ballistic processes with manual movements (e.g., De Jong et al., 1990; Logan, 1981) and arm movements (McCrary and Franks, 1997; McCrary et al., 2000), but their conclusions extend to the control of other effectors, such as control of eye movements (e.g., Boucher et al., 2007a; Kornyllo et al., 2003).

3. Independence of the go and stop process

The independent horse-race model assumes independence between the finishing times of the go process and stop process
The independence assumption takes two forms: context independence (also referred to as signal independence) and stochastic independence. Context independence refers to the assumption that the go RT distribution is the same for no-stop-signal trials and stop-signal trials. Stochastic independence refers to the assumption that trial-by-trial variability in go RT is unrelated to trial-by-trial variability in SSRT \([i.e., \ p(go \ RT < t \land \ SSRT < t) = p(go \ RT < t) \times p(SSRT < t)]\). The independent horse-race model makes these assumptions to simplify the formal model (see Logan and Cowan, 1984). However, the independence assumptions should not be taken lightly because interpretation of the inhibition function and SSRT estimates depend on the validity of the formal model.

3.1. Testing the independence assumptions

The independent horse-race model makes specific predictions about RTs for trials that escape inhibition (i.e., signal-respond RTs). The independence assumptions can be tested by analyzing mean signal-respond RT and by analyzing RT distributions for signal-respond trials.

3.1.1. Tests of mean signal-respond RT

The context independence assumption can be tested qualitatively by comparing mean signal-respond RT with mean no-stop-signal RT, and quantitatively by comparing observed mean signal-respond RT with mean signal-respond RT predicted by the independent horse-race model.

First, the independent horse-race model predicts that mean no-stop-signal RT should be longer than mean signal-respond RT. Mean no-stop-signal RT represents the mean of all responses (i.e., including the longer tail of the go RT distribution) whereas mean signal-respond RT represents the mean of those responses that were fast enough to finish before the stop signal (i.e., excluding the longer tail of the go RT distribution; see Logan and Cowan, 1984; Osman et al., 1986). This can be seen in Fig. 2B: mean no-stop-signal RT represents the mean of the whole go RT distribution whereas the mean signal-respond RT represents the mean of the proportion of the go RT distribution that is on the left of the dashed line (i.e., to the left of the point at which RTir occurs). Several stop-signal studies showed this RT difference between signal-respond and no-stop-signal RTs in different groups, situations, tasks, and conditions (e.g., Aron and Poldrack, 2006; De Jong et al., 1990; Hanes and Schall, 1995; Logan et al., 1984; Osman et al., 1986; Stahl and Gibbons, 2007; van Boxtel et al., 2001; van den Wildenberg and van der Molen, 2004; Verbruggen et al., 2004). When the fixed-SSDs procedure is used, mean signal-respond RTs for the different SSDs can be compared. The independent horse-race model predicts that mean signal-respond RT should increase when SSD increases because more of the go RTs will finish before SSD + SSRT \((\text{Logan and Cowan, 1984; Osman et al., 1986})\). At short SSDs, only the fastest go RTs will escape inhibition. At longer SSDs, slower go RTs may also escape inhibition. This can be seen in Fig. 3A. Consequently, mean signal-respond RT should increase when SSD increases. This prediction has been confirmed by several studies (see e.g., De Jong et al., 1990; Hanes and Schall, 1995; Logan et al., 1984; Osman et al., 1986). However, this analysis requires a fairly large number of trials for every SSD; this may also explain why some studies failed to find that mean signal-respond RT increased when SSD increased (see e.g., Logan, 1981). Indeed, this prediction is more likely to be violated at the shortest SSDs (i.e., with the smallest number of signal-respond trials) than at the longest SSDs (i.e., with the largest number of signal-respond trials; see e.g., Logan et al., 1984).

Second, mean signal-respond RT can be predicted based on the independent horse-race model. Given the observed go RT distribution for no-stop-signal trials and \(p(\text{respond}|\text{signal})\) for a certain SSD, mean signal-respond RT can be predicted by rank-ordering RTs and calculate the mean of the \(n\) fastest no-stop-signal RTs, where \(n\) is obtained by multiplying the number of RTs in the distribution by the probability of responding at a given delay (Logan and Cowan, 1984). The independence assumption is then tested by comparing observed signal-respond RT with predicted signal-respond RT. Some studies reported small differences between observed and predicted signal-respond RT (e.g., De Jong et al., 1990; Hanes and Schall, 1995; Logan and Cowan, 1984), which suggests context independence between the go process and the stop process. Other studies reported larger significant differences between the observed mean signal-respond RT and the predicted mean signal-respond RT (e.g., Colonius et al., 2001; van Boxtel et al., 2001; van den Wildenberg et al., 2002; Verbruggen et al., 2004). In studies that used fixed SSDs, the differences were observed particularly at the shortest stop-signal delays (see e.g., Colonius et al., 2001). The latter findings suggest that the context independence assumption of the horse-race model is sometimes violated.

However, two factors mitigate this conclusion. First, the method for generating the predictions assumes that SSRT is a constant, and that assumption is likely to be false. The assumption of constant SSRT may not bias most estimates of SSRT (see above), but it excludes trials on which go RT is longer than SSD + mean SSRT from the calculation of predicted signal-respond RT. Most likely, this will reduce the predicted mean signal-respond RT: the go process may win the race against the stop process when go RT is longer than SSD + mean SSRT because the latency of the stop process on that trial happens to be longer than the mean SSRT (i.e., go RT > SSD + SSRTM, but go RT < SSD + SSRTn; SSRTM = the mean latency of the stop process and SSRTn = the latency of the stop process on trial \(n\)). Thus, longer signal-signal RTs may be excluded from the calculation of predicted signal-respond RT. Second, simulations performed by Band et al. (2003) suggest that the difference between observed and predicted signal-respond RT is strongly influenced by variability in SSRT, and to a lesser degree, by variability in go RT. The difference between observed and predicted signal-respond RTs increased when variability in SSRT increased, even when the go and stop process were completely independent (variability in go RT had the opposite effect). These findings suggest that a significant difference between the observed and predicted signal-respond RT does not necessarily implies that the independence assumptions of the horse-race model are violated. Note that the comparison between observed and predicted signal-respond RTs does not provide a strong test of the stochastic independence assumption either. Band et al. (2003) showed that manipulating stochastic dependence between the go and stop process did not alter the difference between observed and predicted stop-signal RT much, compared with the effects of variability in SSRT and go RT.

In sum, comparisons of mean signal-respond RT and mean no-stop-signal RT can be a valuable qualitative test of the assumptions of the independent horse-race model. When the predicted differences are not found (or when the opposite difference is found; i.e., mean signal-respond RT is longer than mean no-stop-signal RT), the context independence assumption of the horse-race model may be violated and this may have consequences for the estimation and interpretation of SSRT (see below). By contrast, the difference between predicted signal-respond RT (based on the \(n\) fastest go RTs) and observed signal-respond RT do not provide a strong quantitative test of the context independence assumption of the independent horse-race model.
3.1.2. Tests of signal-respond RT distributions

The context independence assumption can be tested qualitatively by comparing cumulative RT distributions for signal-respond and no-stop-signal trials, and quantitatively by fitting the independent horse-race model to signal-respond RT distributions.

First, the context independence assumption can be tested qualitatively by comparing RT distributions for signal-respond and no-stop-signal trials. Osman et al. (1986) predicted an ordering of cumulative RT distributions for signal-respond and no-signal trials. They predicted that signal-respond and no-stop-signal distributions would have a common minimum and diverge at longer SSDs, with the signal-respond distribution to the left of the no-stop-signal distribution (see Fig. 5). Moreover, signal-respond distributions for shorter SSDs would be to the left of signal-respond distributions for longer SSDs. These predictions stem from the same assumptions that led to the predictions for mean signal-respond RTs: when SSD is short, only the fastest responses will finish before the stop process. When SSD increases, slower responses will also finish before the stop process. Finally, for no-stop-signal trials, all responses, including the slowest ones, will contribute to the RT distribution. Consequently, the distributions will fan out. However, the fastest go RTs contribute to all RT distributions, so the minimum of all distributions should be the same (see Fig. 5). Several studies showed this difference between RT distributions (e.g., Boucher et al., 2007a; Camalier et al., 2007; Osman et al., 1986, 1990), supporting the context independence assumption of the independent horse-race model.

Second, the independence assumptions can be tested quantitatively by fitting the independent horse-race model to signal-respond and no-stop-signal RT distributions. Boucher et al. (2007a) fit the independent horse-race model to these distributions using the assumption that SSRT and go RT are both variable. They modeled the processes underlying the SSRT distribution and go RT distribution (see below), and found that the independent horse-race architecture predicted the signal-respond RT distribution and no-stop-signal RT distribution very well. Camalier et al. (2007) used a different approach, focusing on the finishing-time distributions rather than the underlying processes that generated them. They modeled the independent horse race by sampling finishing times for the go process and the stop process from independent Weibull distributions and found that the model predicted the signal-respond RT and no-stop-signal RT distributions very well. Combined, these fits suggest that the independent horse-race model can predict signal-respond RT very well. However, modeling may not be suited very well for testing the independence assumptions in experiments with small numbers of observations per subject. The qualitative tests may be more practical for many experiments.

3.1.3. Do the independence assumptions really matter in practice?

Logan and Cowan (1984) assumed stochastic and context independence to simplify the formal horse-race model. SSRT estimations are based on the formal independent horse-race model, so one obvious question is what the consequences of violations of the independence assumptions might be. In other words, does independence between the go and stop process really matter in practice?

De Jong et al. (1990) performed a series of Monte Carlo simulations to examine the effect of violations of the independence assumptions. They showed that SSRTs that were estimated by the integration method were less reliable when the go and stop process were correlated (when $r > .2$). This was true mainly for the non-central SSDs (i.e., SSDs for which $p(\text{signal})$ was substantially lower of higher than .50). Band et al. (2003) extended the work of De Jong et al. (1990) by performing a more systematic series of simulations to test how violations of the independence assumptions and variability in the go process and stop process influenced inhibition functions and SSRT. They showed that the central SSRT estimates (i.e., SSRT estimations based on the mean or median of the inhibition function) were relatively unaffected by ‘minor’ violations of the independence assumptions. By contrast, the integration method did not yield reliable SSRT estimates when the independence assumptions were violated (especially because of the non-central SSDs, which is consistent with earlier findings of De Jong et al., 1990). This limitation on the integration method stems from its assumption that SSRT is constant. In order to simulate dependence between SSRT and go RT, some variability has to be introduced into the simulated SSRTs, violating the assumption that SSRT is constant. Thus, the integration method may be more susceptible to violations of the independence assumptions than the methods that assume that SSRT is a random variable. Note that Band et al. also showed that stochastic dependence between the stop and the go process influenced the slope of the inhibition function (even after a ZRTF transformation). Again, this finding suggests that differences in the slope of the inhibition function should be interpreted with caution.

Thus, the answer to the question raised at the beginning of this section is “Yes, the independence assumptions do matter in practice.” Both SSRT estimates and the slope of the inhibition functions can be influenced by violations of the independence assumptions. There are no strong tests available to test minor violations of the independence assumptions, so researchers should choose their SSRT estimation method carefully. The simulation results suggest that the mean method and the median method are relatively unaffected by minor violations of the independence assumptions. The integration method is affected by violations of the independence assumptions, but mainly for non-central SSDs. Thus, the mean method and the median method will result in the most reliable SSRT estimates when fixed delays are used. The integration method should be used primarily with central SSDs. When the tracking procedure is used, both the mean method and the integration method should produce reliable SSRT estimates because one central SSD is used. However, the reliability of the integration method for the tracking procedure has not been explicitly tested yet.

3.2. A neural paradox: the interactive horse race

The independent horse-race model of Logan and Cowan (1984) captures most aspects of go and stop performance. This poses an
interesting paradox: how can a model that assumes complete independence between the finishing times of the go and stop process fit the data so well when the many findings in the cognitive neuroscience literature suggest there are strong interactions between the go process and the stop process. For example, there is overwhelming evidence that eye movements are generated through a network of mutually inhibitory gaze-shifting and gaze-holding neurons (for a review, see e.g., Schall et al., 2002). Yet, the independent horse-race model of Logan and Cowan (1984) accounts for inhibition of eye movements very well (see e.g., Hanes and Schall, 1995).

To address this neural paradox, Boucher et al. (2007a) proposed a variant to the independent horse-race model: the interactive horse-race model. Like the independent horse-race model, the interactive model assumes a race between a go process and a stop process. The go process is initiated by the presentation of the go stimulus and a go unit is activated (viz., movement-related neurons) after an aff erent delay (and possibly, a central decision delay). Similarly, the stop process is initiated by the presentation of the stop signal and a stop unit is activated (viz., fixation-related neurons) after an aff erent delay. When the stop unit is activated, it imposes strong inhibition on the go unit. If that inhibition reaches the go unit before go activation reaches threshold, the go activation may be suppressed enough to prevent it from reaching threshold, resulting in a signal-inhibit trial. If inhibition from the stop unit reaches the go unit too late to prevent go activation from reaching threshold, the go response is executed, resulting in a signal-respond trial.

The interactive horse-race model assumes that the go process and stop process are independent during the initial delay period of the stop process, but they interact when the stop unit becomes active. Tests of the model showed that it fit the behavioral data if and only if the initial delay period of the stop unit was relatively long and the inhibition of the stop unit on the go unit was very strong and thus very brief (whereas the inhibition of the go unit on the stop unit was weak). Thus, the interactive horse-race model fit the data if and only if the stop and go processes were independent for most of their durations, approximating the independence assumptions of the independent horse-race model. Moreover, these fits showed that SSRT primarily reflects the period before the stop unit is activated (i.e., the initiation stage), during which stop and go processing are independent. Thus, SSRT estimates from the independent horse-race model are good behavioral estimates of the duration of the stop process in the interactive horse-race model. More generally, all the behavioral predictions of the interactive horse-race model (including predictions about inhibition functions and signal-respond RT) approximate the behavioral predictions of the independent horse-race model.

A major purpose for developing the interactive horse-race model was to account for neurophysiological data gathered from frontal eye fields in monkeys performing a stop-signal task with eye movements (see e.g., Hanes et al., 1998). Fits of a particular instantiation of the independent horse-race model and a similar instantiation of the interactive horse-race model showed that the two models accounted for the monkeys’ behavior equally well. However, the interactive horse-race model also accounted for the modulation of activity in movement-related and fixation-related neurons in frontal eye fields on stop-signal trials. Boucher et al. used the parameters that provided the best fit to the behavioral data to plot the time-course of activation of the model’s stop and go units, and found that stop and go units modulated on stop-signal trials just like in the neurons in the frontal eye fields. Estimates of the time between the modulation of go unit and SSRT from the model simulation were indistinguishable from estimates of the time between movement-cell modulation and SSRT from the monkeys’ behavior.

In sum, the interactive race model accounts for both behavioral and neurophysiological data, and provides a detailed description of the go and stop process in a stop-signal task with eye-movements. Moreover, it provides a good account of how the idea of an independent race between the go and stop processes can be reconciled with the observation of strong interactions between stop and go processes at a neural level. However, the independent horse-race model has greater generality: it applies to any situation in which there is a meaningful distribution of finishing times. It applies to discrete actions as well as ongoing actions, such as movement tracking (e.g., Morein-Zamir et al., 2004), typing (Logan, 1982) or speech (e.g., Slevc and Ferreira, 2006; Xue et al., 2008). By contrast, the interactive horse-race model applies to the onset of movements and not to ongoing movements, and it is currently very specific to the inhibition of eye movements.

3.3. Functional dependence between the go and stop process

Neurophysiological data argue against a complete independence between the stop and the go process. Several behavioral studies also showed that the go and stop process may be functionally dependent. Logan et al. (1984) showed that the latency of the stop process increased when the go task involved response selection (see also e.g., Serences et al., in press). They showed that SSRTs were longer for a choice RT task, which involved response selection (i.e., every stimulus required a different response), than for a simple RT task, which involved no response selection (i.e., every stimulus required the same response). This finding suggests that the stop process and primary-task processes, such as response selection (or processes that accompany response selection, such as error monitoring), are not completely independent functionally.

Several behavioral studies also showed a functional relation between stop-signal inhibition and interference control in tasks such as the Stroop task, the Eriksen flanker task and the Simon task (e.g., Chambers et al., 2007; Kramer et al., 1994; Ridderinkhof et al., 1999; Verbruggen et al., 2004, 2005a, 2006). These studies showed that SSRT was longer for incongruent trials than for congruent trials. Other studies showed that SSRT was not influenced by resolving interference caused by spatially incompatible responses (e.g., a left-handed response for a rightward pointing arrow; Logan, 1981; van den Wildenberg and van der Molen, 2004), interference caused by switching between tasks (Verbruggen et al., 2005b), or interference due to ignoring the target on the previous trial (i.e., negative priming; Verbruggen et al., 2005c). Combined, these findings suggest that the stop process and some (but not all) kinds of inhibitory control processes in the primary task are functionally dependent. Possibly, the same inhibitory mechanism is involved in stopping and certain types of interference control. This idea is consistent with findings in individual-difference studies, which showed correlations between stop-signal inhibition and the kind of interference control that is involved in the Stroop paradigm and the flanker paradigm (Friedman and Miyake, 2004). The observed functional dependence between different kinds of inhibition is also consistent with neuroimaging studies, which showed overlapping brain structures in different inhibitory tasks (Aron et al., 2004; Derrfuss et al., 2004; Wager et al., 2005). Note that this need not imply that the same inhibitory circuit is involved. Repetitive transcranial magnetic stimulation of the right IFG influenced response inhibition but not interference control in a flanker task with stop signals (Chambers et al., 2007; see also Chambers et al., in this special issue). Future research should clarify whether functional dependence between different kinds of inhibition implies similar neural mechanisms.
In sum, several studies suggest that the go process and the stop process may be functionally dependent in certain tasks. However, this does not imply that the independence assumptions of the independent horse-race model are violated. First, functional dependence does not necessarily imply stochastic dependence: the finishing time of the go process and the stop process may be influenced by the same factor but that does not necessarily imply that these finishing times are stochastically dependent (Ridderinkhof et al., 1999). Stochastic independence implies that \( p(A \cap B) = p(A) \times p(B) \). Some factor could increase both \( p(A) \) and \( p(B) \) without affecting the relation between \( p(A \cap B) \) and \( p(A) \times p(B) \). Second, the period of interaction between the go and stop process is most likely very brief (Boucher et al., 2007a; see Fig. 4). The ‘common inhibitory mechanism’ hypothesis predicts that interference control would influence only the activation stage of the stop process and not the initiation stage. This suggests that when stop signals are introduced in tasks such as the flanker task or the Stroop task, go processing (i.e., responding to the target while ignoring the distractor) and the stop processing (i.e., inhibiting activation in the go unit) are functionally independent for most of their duration, just as they are neurally independent for most of their duration. Thus, functional dependence between the stop process and the go process does not violate the independence assumptions of the independent horse-race model.

3.4. How to balance the go and stop process

Successful performance in the stop-signal paradigm involves monitoring of the go process and the stop process, and adjusting response strategies to balance the competing demands of the two processes. Success in the go task implies failure in the stop task and vice versa. Fast go processes result in a high \( p(A) \times p(B) \). Success in the go task implies failure in the stop task and not the initiation stage. This suggests that subjects balance stopping and going by adjusting response thresholds. This idea was further supported by diffusion-model fits, which allowed quantitative estimates of response thresholds. The diffusion-model fits showed that the response threshold was influenced by the preface. They also showed that non-decision parameters (such as the duration of the motor stage) were sometimes influenced by the preface, which led us to suggest that part of the go RT slowing in stop-signal blocks could be due to proactive (tonic) suppression of motor output. Subjects can also proactively adjust response strategies for specific responses. Aron et al. (2007) told subjects to stop one response when a stop signal was presented (i.e., the critical response) but not the other response (i.e., the non-critical response) in a two-choice RT task. They found that RT was longer for critical responses than for non-critical responses on no-stop-signal trials, which suggests that subjects proactively slowed the go process for critical responses (for similar results, see De Jong et al., 1995).

Several studies suggest that subjects also make reactive response-strategy adjustments after stop-signal trials (e.g., Emerich et al., 2007; Li et al., 2006; Rieger and Gauggel, 1999; Schachar et al., 2004; Verbruggen and Logan, 2008b; Verbruggen et al., 2008a). Rieger and Gauggel (1999) found that go RTs for no-stop-signal trials were prolonged when a stop signal was presented on the previous trial. They suggested that subjects change their response strategy after successful and unsuccessful inhibition to increase the probability of stopping on the next trial. Other researchers suggested that response strategies changed only after unsuccessful stopping (e.g., Schachar et al., 2004; Verbruggen and Logan, 2008b; Verbruggen et al., 2008a). Schachar et al. and Verbruggen et al. suggested that subjects interpret responses on stop-signal trials as errors (see also e.g., Li et al., 2006; Rieger and Gauggel, 1999; van Boxtel et al., 2005) and this leads to reactive response-strategy adjustments, which are reminiscent of the common finding that subjects slow down after making errors in choice-response tasks (Rabbitt, 1966, 1968), trading speed for accuracy. Combined, these studies show that subjects make reactive response-strategy adjustments after stop-signal presentation (Rieger and Gauggel, 1999) or after unsuccessful inhibition (Schachar et al., 2004; Verbruggen and Logan, 2008b; Verbruggen et al., 2008a). However, go accuracy does not always increase when subjects make reactive response strategy adjustments (e.g., Verbruggen and Logan, 2008b; see also Verbruggen et al., 2008a, Experiment 2), which suggests that proactive response-strategy adjustments and reactive response-strategy adjustments influence primary-task performance differently.

Recent studies suggest that repetition of the stimulus that occurred on a stop-signal trial may be a critical variable. In several experiments, we observed slowing after successful stopping, but only when the stimulus or stimulus category of the previous trial was repeated. This led us to suggest a memory-retrieval explanation for the after-effects of successful inhibition: the primary-task stimulus or stimulus category is associated with the stopping on a stop-signal trial; when the stimulus (or category) is repeated, the stimulus-stop association is retrieved, and this interferes with go responding on no-stop-signal trials. These effects are observed up to 20 trials after the presentation of the stop signal, which suggests that the stimulus-stop associations are stored into memory in the form of long-term associations (Verbruggen and Logan, 2008b). These long-term associations may support the development of automatic response inhibition, in which inhibition is driven primarily by bottom-up retrieval of stimulus-stop associations instead of top-down activation of the stop process (Verbruggen and Logan, in press-b). Note that automatic inhibition is more likely to develop in the go/no-go paradigm, where stimuli are consistently associated with going and stopping, than in the stop-signal paradigm, where stimuli are
inconsistently associated with going and stopping. Consequently, the two paradigms may put different demands on cognitive control (Verbruggen and Logan, in press-a).

4. The role of the stop process in inhibiting a response

Logan and Cowan (1984) described how respond inhibition depends on the relative finishing time of a go process and a stop process. Boucher et al. (2007a,b) elaborated this idea and described how a stop unit strongly inhibits a go unit after an affrent delay. Central to these models is that a go response is inhibited by the activation of a stop process. An alternative to this idea is that a go response is inhibited by the preparation of an alternative go response. In this case, response inhibition would depend on the relative finishing time of the primary-task response (the go1 response) and the alternative response (the go2 response).

Recently, we tested the ‘alternative response’ hypothesis using the stop-change paradigm (Verbruggen et al., 2008c). The stop-change paradigm is similar to the standard stop-signal paradigm in that subjects are instructed to stop their response for the primary task (hereafter referred to as the go1 task) whenever a stop-change signal is presented. But in addition, subjects have to replace the stopped response with a new response for a secondary task (the go2 task). The go2 task has been implemented in several ways. In some studies, subjects just pressed a key that was not used in the go1 task (Logan and Burkell, 1986) or they pressed the opposite go1 key (e.g., press the left key instead of the right key; e.g., Brown and Braver, 2005; Nachev et al., 2007); in other studies, subjects responded to the identity of the stop-change signal (e.g., discriminating whether the stop-change signal was a high or a low tone; De Jong et al., 1995; Logan, 1983, 1985b; Verbruggen and Logan, 2008a).

We introduced a delay between the stop signal and the go2 signal to test whether the go1 response can be inhibited by activating the go2 response (Verbruggen et al., 2008c). This manipulation allowed us to distinguish between different models that were built around two questions: first, is a stop process involved in inhibiting the go1 response, and more specifically, does inhibition of the go1 response depend on the relative finishing time of a race between the go1 process and the stop process or does it depend on the relative finishing time of the go1 process and the go2 process? Second, if a separate stop process is involved, can the stop process and the go2 process be activated simultaneously? The results of two experiments were consistent with the models that included a stop process. We found that p(respond signal) was barely influenced by the delay between the stop signal and the go2 signal. By contrast, go2 RT decreased substantially when the delay between the stop signal and the go2 signal increased, which is reminiscent of the psychological-refractory period effect (Pashler, 1994). These findings led us to conclude that the go1 response is inhibited by the activation of a stop process and not by the activation of the alternative go2 response. Moreover, the alternative go2 response seemed to be activated after the stop process finished, most likely because of strategic limitations (Logan and Gordon, 2001; Meyer and Kieras, 1997), not structural limitations (i.e., a central-processing bottleneck; Pashler, 1994). Successful response inhibition depends on the relative finishing time of the go1 process and the stop process, so it is beneficial to allocate all processing capacity to stopping even though response inhibition can be selective (i.e., the go1 process can be inhibited without inhibiting the go2 process; see e.g., Aron and Verbruggen, in press; De Jong et al., 1995).

Camalier et al. (2007) compared two oculo-motor procedures that are similar to the stop-change paradigm: the double-step task and the search-step task. In both tasks, subjects are required to make an eye movement to a target. On double-step trials, the location of the target changes before the initial eye movement is made and subjects have to make an eye movement to a new location. On search-step trials, the initial target becomes a distractor and a stimulus that was previously a distractor becomes the new target. Thus, in both tasks subjects have to stop and change an eye movement that is no longer relevant. Camalier et al. (2007) distinguished between three models of task performance built around the same questions that were addressed by Verbruggen et al. (2008c). To determine whether a stop process was necessary, they compared models that assumed an explicit stop process with a model that assumed a race between the go1 process and the go2 process. To determine whether the go2 process could begin before the stop process finished, they compared a serial stop model that assumed that go2 processing started when the stop process finished with a parallel stop model that assumed that go2 processing started at the same time as the stop process. The three models were fit to the data of both humans and macaque monkeys. The model fits suggested that a stop process was necessary: the models that included a stop process fitted the data better than the model without it. The serial and parallel stop models fitted the data equally well. Camalier et al. had no experimental manipulation that allowed them to discriminate between the two stop models, so they distinguished between them on grounds of plausibility. The serial model fits produced go2 RTs that were go1 RTs. This seemed implausible, so Camalier et al. (2007) favored the parallel model.

In sum, recent work with the stop-change paradigm suggests that response inhibition requires a separate stop process. Subjects cannot stop and replace a response by simply activating an alternative response. A stop process must inhibit the go1 response before the go2 response can be executed. Stop performance benefits from a response inhibition mechanism because it operates faster than selecting and preparing an alternative response.

5. Concluding remarks

The stop-signal paradigm is very useful for the study of response inhibition in a laboratory setting judging from the widespread use of the paradigm in cognitive psychology, clinical psychology, cognitive neuroscience and neuropsychology. Performance in the stop-signal paradigm is typically described as a race between a go process and a stop process. This horse race was formalized by Logan and Cowan (1984). After more than 25 years, their independent horse-race model still offers a remarkably good account for stop-signal performance in a variety of settings, across different populations, tasks and conditions. On the basis of a few simple assumptions, the independent horse-race model can describe both observable (go RT, p(respond signal)) and unobservable (SSRT) measures of stop-signal performance. Importantly, the model provides theoretically justified estimates of latency of the stop process (SSRT). Recent variants of the independent-horse model account for brief moments of interactions between neurons (i.e., the interactive horse-race model) or for more sustained interactions between multiple go processes (i.e., models for performance in the stop-change paradigm or the double-step and search-step paradigms). These models offer a more detailed description of performance in specific situations but lack the generality of the independent horse-race model. Nevertheless, we believe that this is a fruitful avenue for future research and that the general independent horse-race model can serve as a common basis for more detailed descriptions of performance in a broad range of specific situations.

Acknowledgments

Frederick Verbruggen is a Postdoctoral Fellow of the Research Foundation—Flanders. Gordon D. Logan is supported by grant BCS
0646588 from the National Science Foundation, grant no. FA9550-07-1-0192 from the Air Force Office of Scientific Research grant, and by National Institute of Mental Health Grant R01-MH073879-01. We thank Guido Band, Adam Aron and an anonymous reviewer for their useful comments on an earlier version of this manuscript.

References


Morein-Zamir, S., Meiran, N., 2003. Individual stopping times and cognitive control: converging evidence for the stop signal task from a continuous tracking para-


