Modeling response inhibition in the stop signal task

Hans Colonius Adele Diederich

This is a chapter for volume 3 of the *New Handbook of Mathematical Psychology*, edited by F.G. Ashby, H. Colonius, and E.N. Dzhafarov (Cambridge UP, 2023)

Contents

0.1 Response inhibition and the stop signal task

The notion of *response inhibition* refers to an organism's ability to suppress unwanted impulses, or actions and responses that are no longer required or have become inappropriate. This ability is considered a case of cognitive control, those cognitive faculties that allow information processing and behavior to vary adaptively from moment to moment depending on current goals, rather than remaining rigid and inflexible. At this time, the field of cognitive control flourishes like never before (Logan, 2017, p. 875).[1]

The simple fact that cognitive control takes time makes subjects' behavior amenable to the advanced methods of response time analysis and modeling developed in cognitive psychology over many years. In the *stop-signal paradigm*, participants typically perform a *go task* (e.g. press left when an arrow pointing to the left appears, and press right when an arrow pointing to the right appears), but on a minority of the trials, a stop signal (e.g. an acoustic stimulus) appears after a variable stop-signal delay, instructing the participant to suppress the imminent go response (see Figure 0.1). This paradigm has become the main workhorse being used in laboratory settings across various human populations (e.g. clinical vs. non-clinical, different age groups) as well as non-human ones (primates, rodents, etc.).

The stop-signal task provides three types of observable data: (i) reaction times (RTs) to the go signal in go trials, (ii) RTs in stop trials (when response inhibition failed), and (iii) the frequency of responses given in spite of the stop signal. Unlike the latency of go responses, response-inhibition latency cannot be observed directly (as successful response inhibition results in the absence of an observable response). This is a problem, in particular because the time to cancel a response is widely considered to be an appropriate indicator of the level of response inhibition of an individual, and it must be addressed by any model of the stop-signal task.

[1] Gordon Logan emphasizes that "Cognitive control addresses core issues in basic and applied psychology, from free will and the nature of intention to practical strategies for improving our own control and treating deficient control in our clients.". According to *Web of Science* (10/2020), there were about 200 papers and 10k citations in 2019 for "stop-signal task".

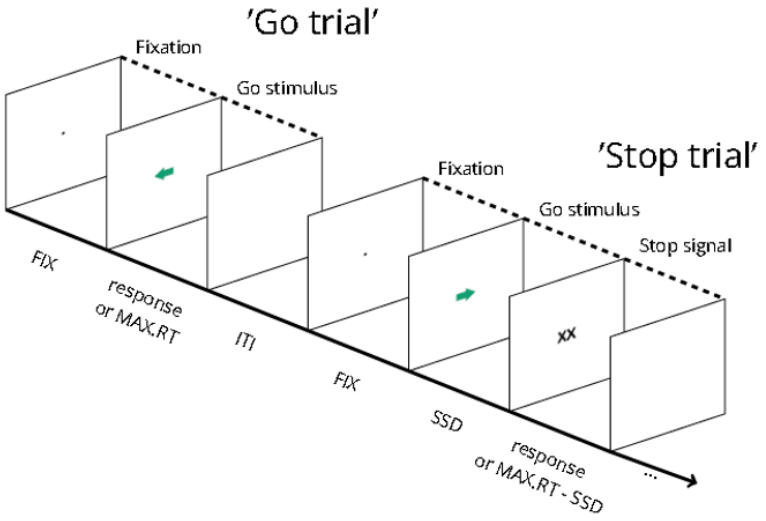

Figure 0.1 Depiction of the sequence of events in a stop-signal task. In this example, participants respond to the direction of green arrows (by pressing the corresponding arrow key) in the go task. On one fourth of the trials, the arrow is replaced by "XX" after a variable stop-signal delay (FIX = fixation duration; SSD = stop-signal delay; MAX.RT = maximum reaction time; ITI = intertrial interval) (from Verbruggen et al., 2019).

The main goal of this chapter is to present results in the formal modeling of behavioral data from the stop-signal paradigm and some of its variants. Given that there exists a number of comprehensive literature reviews of both empirical and modeling results (see final section on bibliographic notes), we primarily present a general formal framework allowing us to incorporate most current models and, at the same time, expose a number of open or only partially solved problems. In order to keep the chapter self-contained, we start by presenting some typical data pattern. Then, the general race model is introduced including estimation methods for the non-parametric case (Section 0.3). More detailed presentations of parametric independent (Section 0.4) and dependent (Section 0.5) race models follow. Some related, but non-race, models are discussed in Section 0.6. Section 0.7 introduces the class of semi-parametric race models based on the copula concept. It also contains the race model with perfect negative dependence. Variants of the stop signal paradigm, the problems of trigger failures and sequential effects are sketched in Section 0.8. We conclude with a brief discussion contrasting parametric versus non-parametric approaches and a look into the future of

Acronym	Meaning
ARI	anticipated response inhibition
CI	context independence
FEF	frontal eye field
FGM	Farlie-Gumbel-Morgenstern (copula)
IT	inhibition time
LATER	linear approach to threshold with ergodic rate (model)
MCMC	Markov chain Monte Carlo
PND	perfect negative dependency
PTC	pause-then-cancel (model)
RT	response (or, reaction) time
SC	superior colliculus
SI	saccadic inhibition
SOA	stimulus-onset asynchrony
SSD	stop-signal delay
SSRT	stop-signal reaction time

Table 0.1 *Abbreviations used in the chapter*

stop signal modeling. A list of abbreviations used in the chapter is found in Table 0.1.

0.2 Some typical data patterns in the stop-signal paradigm

Given the popularity of the stop-signal task, the amount of data is enormous and, unsurprisingly, there is a lot of diversity in the findings due to differences in design, instruction, and the specific sub-population tested. Nonetheless, many results only differ with respect to their specific numerical values observed for reaction times and inhibition probabilities, while some general qualitative features of the inhibition function and RT distributions are typically retained.

0.2.1 Inhibitions function

Inhibitions functions depict the probability of a response in spite of a stop signal as a function of stop-signal delay (SSD)[2]. When the stop signal occurs soon after the go signal, participants have a high chance of withholding a response, so the inhibition function has a small value. With SSD increasing, this chance diminishes more and more, up to a point where the probability to respond approaches 1. The top panel of Figure 0.2 depicts classic data

[2] Strictly speaking, it should be called "non-inhibition function", but the terminology used here is common.

from three subjects reported in Logan and Cowan (1984). While these in-

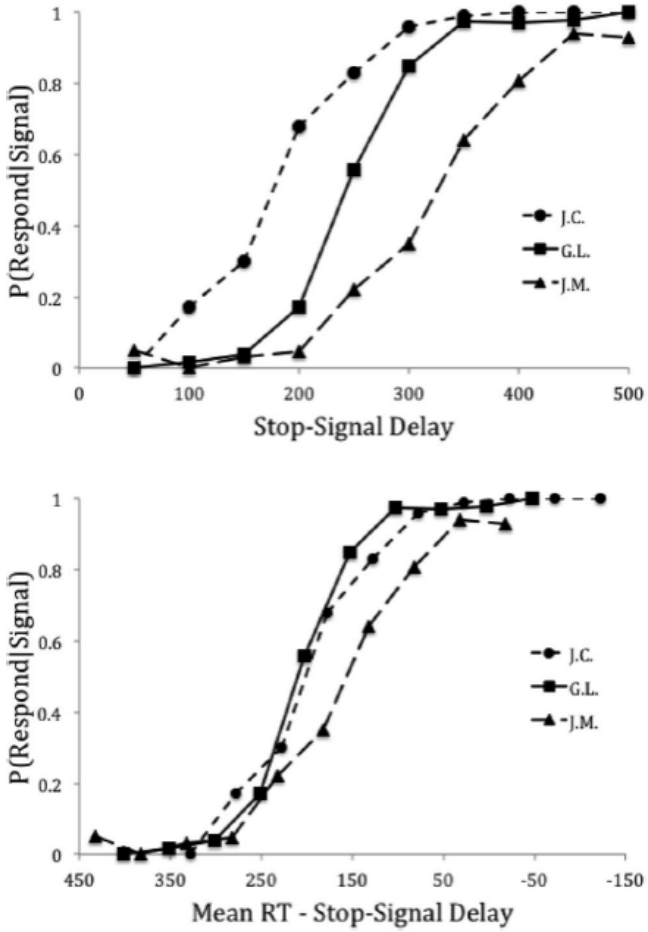

Figure 0.2 Classic data reported in Logan and Cowan (1984). Top panel: Inhibition functions from 3 subjects plotted as function of stop-signal delay. Bottom panel: Inhibition probability for the same three subjects replotted as a function of mean go response time minus stop-signal delay (SSD) (from Logan et al., 2014).

hibition functions are somewhat similar in shape, subjects clearly differ: for mid-range SSD values, the probability of a response can vary enormously. Does this imply that, e.g., participant J.M. (lower curve) is much better in controlling the response than the other two? Unfortunately, interpretation of inhibition functions is not straightforward. Although J.M.'s inhibitory performance may in fact be best, it could also be the result of J.M. voluntarily

slowing responses to the go signal in most trials, so that there is always "enough" time to stop the response. Even if one persuades participants to not delay their response, it has still been shown that various parameters of the distribution of responses to the go signal, like variance, may have a strong effect on the inhibition functions. Suggestions to remove these problems by a standardized transformation of the inhibition function remain controversial, however. The bottom panel of Figure 0.2 shows the probability of inhibition plotted against the difference between mean go RT and stop-signal delay for the same subjects. Under some simplifying assumptions, this difference is interpretable as a measure of the time that is available to detect the stop signal and to cancel a response.[3] In sum, this issue calls for developing a formal model within which the level of performance can be gauged exactly by some parameter estimated from data.

0.2.2 Reaction times to go and to stop signal

The distribution of reaction times on go trials, i.e. without a stop signal, are often more or less right-skewed, as is typical for RT distributions in general. Figure 0.3 depicts the histogram of 2144 saccadic reaction times to a visual target, occurring either to the left or right of the fixation point, in a stop-signal task with auditory stop signals (Özyurt et al., 2003). Responses on

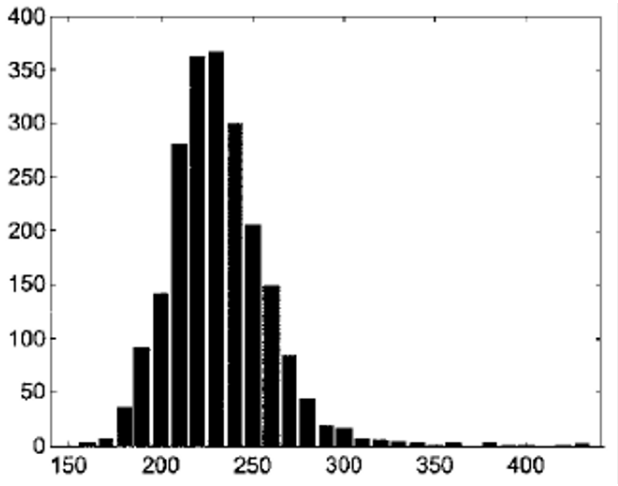

Figure 0.3 Saccadic reaction times to a visual target with $N = 2144$ (from Özyurt et al., 2003, subject P.T.).

[3] Note that the functions for J.C. and G.L. align better than the function for J.M. because J.M. had greater variability in go RT than the other two.

unsuccessful stop trials (*signal-respond* RT) are on average faster than go RT on trials with no stop signal and faster for shorter stop-signal delays than for longer ones. Note that this latter observation is to be expected assuming that the process of inhibition evolves over a possibly variable time interval. This feature (often called "fan effect"), illustrated in Figure 0.4 by another study on saccadic RTs to a visual target with an auditory stop signal (Colonius et al., 2001), motivated the development of so-called race models to be discussed below.

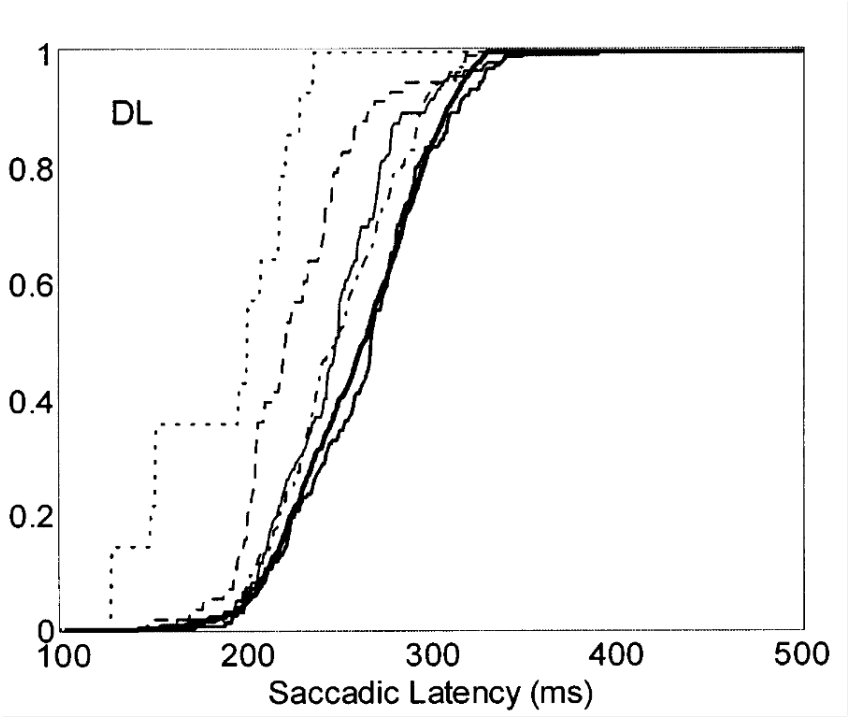

Figure 0.4 Empirical (signal-respond) distribution functions of saccadic RTs to a visual target with an auditory stop signal presented at different SSD values [ms] (in parentheses: number of observations): dotted line: 150 (14); dashes: 200 (58); thin line: 230 (122); dots/dashes: 250 (147); medium line: 270 (170); thick line: go condition (2919) (from Colonius et al., 2001, subject D.L.).

0.3 Modeling the stop signal task

Many features of typical data in the stop-signal task are consistent with modeling responses as the outcome of a race between processing of the go

signal and the stop signal: if the latter terminates earlier than the former, subjects succeed in inhibiting a response, otherwise they respond in spite of the stop signal. Although "race" is the predominant modeling approach, let us first take a step back and consider the situation from a more general point of view.

When only the go signal is presented, denoted as *context GO*, reaction time T_{go}, say, represents the time to process that signal, including possible pre- and motor components. In order to account for some variability across trials, T_{go} is considered a random variable taking on non-negative values. In contrast, when the go signal is followed by presentation of a stop signal, denoted as *context STOP*, the two alternative outcomes –either a response is given or there is no response– are the result of (somehow) processing both signals. Race models hold that, in addition to T_{go}, there is a separate random processing time for the stop signal, T_{stop}, say, and the outcome is determined by $\min\{T_{go}, T_{stop} + \text{SSD}\}$. Alternatively, instead of claiming a separate processing time T_{stop}, one could assume that the stop signal modulates processing of the go signal in a way that is qualitatively consistent with two fundamental empirical observations. First, RTs in context *STOP* tend to be faster than in context *GO*; thus, according to this alternative view, the stop signal speeds up processing time T_{go} for the go signal. Second, the probability of inhibition decreases with SSD; thus, the later the stop signal is presented, the shorter the time it can modify processing time T_{go}. We will sketch such non-race models in Section 0.6.

Nonetheless, a glance over the stop-signal literature strongly suggests that the race model is the "main game in town", especially when certain generalizations and extensions of the notion of "race" are included, like interdependent processing or certain across-trial strategies for optimizing responses. Therefore, the chapter's focus is on this model class. The next section provides a formal introduction of the race model and its subclasses.

0.3.1 The general race model

One important distinction in classifying race models is whether they are parametric or non-parametric, that is, if specific distributional assumptions concerning T_{go} and T_{stop} are made. Another is whether these random variables are considered to be statistically independent or not (see Figure 0.5). Although semi-parametric race models actually contain both parametric and non-parametric instances, they are listed here as a separate subclass for conceptual reasons. They are based on the definition of a copula and will be discussed in Section 0.7.

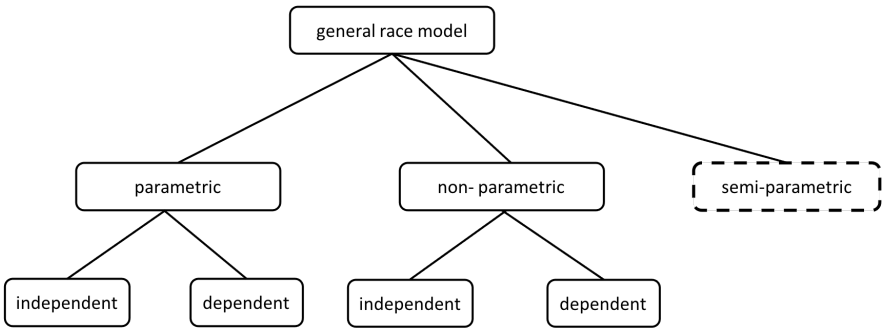

Figure 0.5 Parametric, non-parametric, and semi-parametric subclasses of the general race model with either independent or dependent T_{go} and T_{stop} processing times.

For context $STOP$, we postulate a bivariate cumulative distribution function (cdf), denoted H, for T_{go} and T_{stop},

$$H(s,t) = \mathrm{P}\left[T_{go} \leq s, T_{stop} \leq t\right], \tag{0.1}$$

defined for all real numbers s and t, with $s, t \geq 0$. Moreover, T_{go} and T_{stop} are assumed to be continuous random variables[4]. Sometimes T_{stop} is referred to as *stop signal reaction time* (SSRT). The marginal cdfs of $H(s,t)$ are denoted as

$$\begin{aligned} F_{go}(s) &= \mathrm{P}\left[T_{go} \leq s, T_{stop} < \infty\right] \text{ and} \\ F_{stop}(t) &= \mathrm{P}\left[T_{go} < \infty, T_{stop} \leq t\right]. \end{aligned}$$

In context $STOP$, the go signal triggers realization of random variable T_{go} and the stop signal triggers a realization of random variable T_{stop}. In context GO, however, only processing of the go signal occurs. Thus, the two different experimental conditions in the paradigm, GO and $STOP$, imply the existence of two different sample spaces in the statistical modeling of the task. In principal, the distribution in context GO, $F^*_{go}(s)$, say, could be different from the marginal distribution $F_{go}(s)$ in context $STOP$.

However, the general race model rules this out by adding the important assumption of *context independence*, also known as *context invariance*[5]:

Context independence (CI) In context GO, the distribution of go signal processing time is assumed to be

$$F^*_{go}(s) \equiv F_{go}(s) = \mathrm{P}\left[T_{go} \leq s, T_{stop} < \infty\right] \tag{0.2}$$

[4] That is, H possesses a bivariate density.
[5] *Context invariance* seems a more fitting term but to avoid confusion, we keep the familiar *context independence*.

for all s, i.e., it is identical to the marginal distribution $F_{go}(s)$ in context $STOP$.

Note that, in order to be more precise, context $STOP$ would have to be indexed by the specific value of SSD, t_d, say, with $t_d \geq 0$, and the same holds for $H(s,t)$ and $F_{stop}(t)$. In the following, however, we will tacitly assume that SSD *invariance* holds, meaning that we can drop the index t_d throughout without consequences while keeping it as a given (design) parameter. Moreover, T_{stop} is set equal to zero for $t \leq t_d$, with probability one.

From these assumptions, the probability of observing a response (r) to the go signal given a stop signal was presented with SSD $= t_d$ [ms] after the go signal, is defined by the *race assumption* ,

$$p_r(t_d) = \mathrm{P}\left[T_{go} < T_{stop} + t_d\right]. \tag{0.3}$$

In addition, according to the model, the probability of observing a response to the go signal no later than time t, given the stop signal was presented with delay t_d, is given by the (conditional) distribution function

$$F_{sr}(t \,|\, t_d) = \mathrm{P}\left[T_{go} \leq t \,|\, T_{go} < T_{stop} + t_d\right], \tag{0.4}$$

also known as *signal-respond RT (sr)* distribution.

The main interest in modeling the race is to derive information about the distribution of the non-observable stop signal processing time, T_{stop}, or about some of its parameters given sample estimates of $F_{go}(t)$, $F_{sr}(t \,|\, t_d)$, and $p_r(t_d)$. For example, the independent race model presented in Section 0.3.2 is parameter-free, i.e., no parameters have to be estimated in order to make predictions. Later, we will discuss both fully parameterized models and semi-parametric versions. In the latter, no specific distributions are postulated but only a parameter assessing the degree of stochastic dependency.

The most simple version of the race model, sometimes referred to as *independent horse race model*, assumes the non-observable time $T_{stop} =$ SSRT to be a constant k, $k \geq 0$. Thus, $p_r(t_d)$ becomes simply

$$p_r(t_d) = \mathrm{P}\left[T_{go} \leq t_d + k\right].$$

Figure 0.5 is the standard depiction of this specific model. It illustrates how the probability to respond given a stop signal (the area under the curve to the left of the vertical line) depends on (i) SSD (thus generating the inhibition function), (ii) the go RT distribution, and (iii) the stop signal processing time (SSRT).

Assuming constant stop signal processing time is not realistic and may impair model predictions (Verbruggen and Logan, 2009), but it simplifies

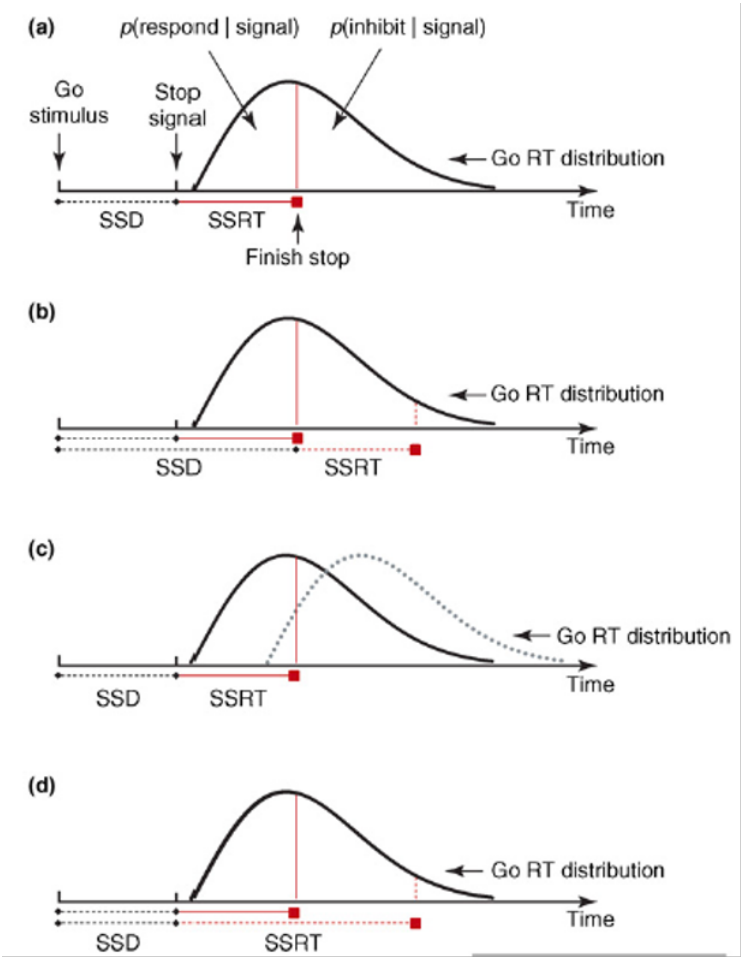

Figure 0.6 Schematic of the simplified race model: the probability to respond given a stop signal (the area under the curve to the left of the vertical line) depends on t_d = SSD (panel b), on go RT distribution (panel c), and stop signal processing time (SSRT = k) (panel d) (from Verbruggen and Logan, 2008).

estimation of SSRT enormously. In fact, a popular estimation method for SSRT, the *integration method*, requires it (see below).

0.3.2 The (complete) independent race model

The most common version of the race model is the (complete) independent race model[6] introduced by Logan and Cowan (1984); it postulates stochastic independence between T_{go} and T_{stop}:

Stochastic independence:

$$H(s,t) = \mathrm{P}\left[T_{go} \leq s\right] \times \mathrm{P}\left[T_{stop} \leq t\right] = F_{go}(s) \times F_{stop}(t), \tag{0.5}$$

for all s, t $(s, t \geq 0)$.

From this, we have

$$\begin{aligned} p_r(t_d) &= \mathrm{P}\left[T_{go} < T_{stop} + t_d\right] \\ &= \int_0^\infty f_{go}(t)[1 - F_{stop}(t - t_d)]\, dt, \end{aligned} \tag{0.6}$$

with $f_{go}(t)$ denoting the probability density function (pdf) for T_{go}. Moreover, the signal-respond distribution equals,

$$\begin{aligned} F_{sr}(t \,|\, t_d) &= \mathrm{P}\left[T_{go} \leq t \,|\, T_{go} < T_{stop} + t_d\right], \\ &= \frac{1}{p_r(t_d)} \int_0^t f_{go}(t')[1 - F_{stop}(t' - t_d)]\, dt'. \end{aligned} \tag{0.7}$$

for all $t > t_d$ and $p_r(t_d) > 0$[7].

The predominance of the independent model is due to the fact that its predictions are mostly consistent with the empirical observations presented above. First, increasing t_d in Equation 0.6 monotonically increases the expression under the integral, thus increasing the probability of a response and approaching 1 in the limit for $t_d \to +\infty$, as observed in Figure 0.2 (top panel). Second, letting $t_d \to +\infty$ in Equation 0.7 implies $F_{sr}(t \,|\, t_d)$ to approach $F_{go}(t)$, for any fixed t (Figure 0.4). As an additional test, the signal-respond distribution has been shown to have an upper and a lower bound (Colonius et al., 2001),

$$F_{go}(t) \leq F_{sr}(t \,|\, t_d) \leq F_{go}(t)/p_r(t_d) \tag{0.8}$$

[6] The attribute "complete" is sometimes used to distinguish this model from the one with constant SSRT.

[7] One can define $F_{sr}(t \,|\, t_d)$ for $t \leq t_d$ as well: Equation 0.7 then results in

$$F_{sr}(t \,|\, t_d) = \min\left\{\frac{F_{go}(t)}{p_r(t_d)}, 1\right\}.$$

It is the probability of an anticipatory response (given even before the stop signal is presented), but these responses are usually removed.

for all t. The lower bound implies, in particular, that

$$\mathrm{E}[T_{go} \,|\, T_{go} < T_{stop} + t_d] \leq \mathrm{E}[T_{go}],$$

i.e. mean stop failure responses should be faster than mean go signal responses.

Writing $f_{sr}(t \,|\, t_d)$ for the pdf of $F_{sr}(t|t_d)$, it follows (Colonius, 1990) that

$$f_{sr}(t \,|\, t_d) = f_{go}(t)\,[1 - F_{stop}(t - t_d)]/p_r(t_d). \tag{0.9}$$

From that, an explicit expression for the distribution of unobservable stop signal processing time (T_{stop}) follows after rearrangement:

$$F_{stop}(t - t_d) = 1 - \frac{f_{sr}(t \,|\, t_d)p_r(t_d)}{f_{go}(t)}. \tag{0.10}$$

Unfortunately, simulation studies revealed that gaining reliable estimates for the stop signal distribution using Equation (0.10) requires unrealistically large numbers of observations (Band et al., 2003; Matzke et al., 2013). As long as one is satisfied with obtaining just an estimate of some parameter of the stop signal distribution, like the mean, two common 'non-parametric' methods are available. If the entire distribution is of interest, a parametric model assuming a distributional family, like the ex-Gaussian, is called for. Both alternatives will be discussed.

0.3.3 Non-parametric estimation of stop signal distribution under independence

We first describe the underlying theoretical assumptions of the methods, followed by some practical considerations for their usage.

Mean method. Rewriting the probability of a response given the stop signal is presented at SSD $= t_d$ as

$$p_r(t_d) = \mathrm{P}\,[T_{go} - T_{stop} < t_d],$$

it can be interpreted formally as the cdf of a random variable T_d, say, taking values t_d, see Logan and Cowan (1984) and illustrated by the shape of Figure 0.2, bottom panel). It follows that $T_{go} - T_{stop}$ and T_d are *equal-in-distribution*.[8] In particular, we get

$$\mathrm{E}[T_{stop}] = \mathrm{E}[T_{go}] - \mathrm{E}[T_d], \tag{0.11}$$

[8] This means they have the same distribution but are not (necessarily) defined on the same sample space; see Chapter 1 in Volume 1 (p. 10) for definitions.

for the mean, and

$$\mathrm{Var}[T_{stop}] = \mathrm{Var}[T_d] - \mathrm{Var}[T_{go}] \tag{0.12}$$

for the variance of T_d, the latter following due to stochastic independence of T_{go} and T_{stop}.

Integration method. In contrast to the mean method, here stop signal processing time is taken to be a constant, t_{stop}, say. Thus,

$$F_{stop}(t) = \begin{cases} 0, & \text{if } t < t_d + t_{stop}; \\ 1, & \text{if } t \geq t_d + t_{stop}. \end{cases} \tag{0.13}$$

Inserting in Equation 0.6 yields

$$\begin{aligned} p_r(t_d) &= \int_0^\infty f_{go}(t)[1 - F_{stop}(t - t_d)]\, dt, \\ &= \int_0^{t_d + t_{stop}} f_{go}(t) dt, \\ &= F_{go}(t_d + t_{stop}), \end{aligned} \tag{0.14}$$

and inserting in Equation 0.7 yields

$$\begin{aligned} F_{sr}(t \,|\, t_d) &= \frac{1}{p_r(t_d)} \int_0^t f_{go}(t')[1 - F_{stop}(t' - t_d)]\, dt', \\ &\begin{cases} \frac{F_{go}(t)}{F_{go}(t_d + t_{stop})}, & \text{if } t < t_d + t_{stop}; \\ 1, & \text{if } t \geq t_d + t_{stop}. \end{cases} \end{aligned}$$

The value of t_{stop} is obtained via Equation 0.14 by determining the quantile of the go-signal distribution, $F_{go}^{-1}(t_d + t_{stop})$, and subtracting the corresponding SSD value t_d (see Figure 0.5).

Some practical considerations. Whether the mean or integration method should be used depends in part on the way the stop-signal delays are set. First, one can simply choose a fixed number of SSDs such that the range of the probability of responding, $p_r(t_d)$, is sufficiently covered. The second method adjusts SSDs dynamically using a tracking procedure (mostly, one-up/one-down), as described, e.g., in Matzke et al. (2018)[9]. At convergence,

[9] "At the beginning of the experiment, stop-signal delay 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, stop-signal delay increases (e.g., by 50 ms); when inhibition is unsuccessful, stop-signal delay decreases (e.g., by 50 ms)."

this results in an approximate value of SSD (t_d) such that $p_r(t_d) = 0.5$, but step size should be optimized to avoid slow or no convergence. The tracking method typically results in a sufficiently varied set of SSD values so that $\mathrm{E}[T_{stop}]$ can be estimated easily by subtracting mean SSD from mean RT on go trials corresponding to Equation 0.11, making the mean method the most popular estimation method.

Applying the integration method with a fixed number of SSDs involves rank-ordering the go RTs for each t_d and selecting the n-th go RT where n is the number of go RTs multiplied $p_r(t_d)$. Stop-signal delay is then subtracted to arrive at an estimate of t_{stop} (compare Equation 0.14). Estimates from different stop-signal delays are averaged to arrive at a single estimate for each participant, also when the tracking procedure is being used. Simulation results reported in Verbruggen et al. (2019) suggest that the integration method produces the most reliable and least biased non-parametric SSRT estimates under the condition that go omissions (i.e. go trials on which the participant did not respond before the response deadline) and premature responses on unsuccessful stop trials (i.e. responses executed before the stop signal is presented) should be included in the estimation procedure. Due to numerous recommendations in the literature on how to conduct stop signal experiments (Logan, 1994; Matzke et al., 2018; Verbruggen et al., 2019), applying non-parametric race models has become a more or less routine task.

0.4 Parametric independent race models

One reason for adopting a parametric distributional family for go and stop signal processing times is the desire to obtain additional measures of inhibition performance, like variance or skew, in order to differentiate, for example, between clinical sub-populations. Another motive is trying to reveal the mechanisms that implement going and stopping and to predict effects of experimental manipulations on stop-signal performance in the context of a substantive process model of response inhibition. A selected set of independent parametric models will be considered here.

In principle, assuming some parametric form for the distributions of T_{go} and T_{stop} and inserting them into the equations for the go and stop signal distributions (Equations 0.6 to 0.9) is straightforward, but obtaining closed-form expressions is often not achievable. The signal-respond distribution can

be written as

$$F_{sr}(t \mid t_d; \theta_{go}; \theta_{stop}) = \mathrm{P}\left[T_{go} \leq t \mid T_{go} < T_{stop} + t_d; \theta_{go}; \theta_{stop}\right],$$
$$= \frac{1}{p_r(t_d, \theta_{go}, \theta_{stop})} \int_0^t f_{go}(t' \mid \theta_{go})[1 - F_{stop}(t' - t_d \mid \theta_{stop})]\, dt', \tag{0.15}$$

with θ_{go} and θ_{stop} denoting parameters, or vectors of parameters, for the go and stop signal distribution, respectively.

A number of different estimation methods for parametric models have been developed. Parameter estimation via maximum likelihood requires the likelihood functions for both go and stop signal conditions that can be written as follows.

Let $\{t_g\}_{g=1,\ldots G}$ denote a sample of G response times collected in context GO. The log-likelihood function becomes

$$\log L(\theta_{go} \mid \{t_g\}) = \sum_{g=1}^{G} f_{go}(t_g \mid \theta_{go}). \tag{0.16}$$

For context $STOP$, we must distinguish stop-signal responses and inhibitions: let $\{t_r\}_{r=1,\ldots R}$ denote the signal-respond times for a given SSD $= t_d$. This implies the following log-likelihood function:

$$\log L(\theta_{go}, \theta_{stop} \mid \{t_r\}, t_d) = \sum_{r=1}^{R} f_{go}(t_r \mid \theta_{go})[1 - F_{stop}(t_r - t_d \mid \theta_{stop})]. \tag{0.17}$$

Turning to the inhibitions, let $\{t_i\}_{i=1,\ldots I}$ denote the successful inhibition (stop signal) times. Because the t_i's are not observable, the likelihood of winning at each possible time point must be considered (by integration). For a given SSD $= t_d$, the log-likelihood function is thus given by (Matzke et al., 2013)

$$\log L(\theta_{go}, \theta_{stop} \mid \{t_i\}, t_d) = \sum_{i=1}^{I} \int_{t_d}^{\infty} \{f_{stop}(t_i - t_d \mid \theta_{stop})[1 - F_{go}(t_i \mid \theta_{go})]\}\, dt_i. \tag{0.18}$$

0.4.1 Exponential model.

We start with the exponential model as an illustrative example permitting closed-form predictions. Several, more prominent models will follow including information about suitable parameter estimation methods.

Let T_{go} and T_{stop} follow exponential distributions; the bivariate cdf is

$$\begin{aligned} H(s,t) &= \mathrm{P}\left[T_{go} \le s\right] \times \mathrm{P}\left[T_{stop} \le t\right] \\ &= (1 - \exp[-\lambda_{go}\, s]) \times (1 - \exp[-\lambda_{stop}\, t]), \end{aligned}$$

for all $s, t \ge 0$ with positive real-valued parameters λ_{go} and λ_{stop} . Then

$$\begin{aligned} p_r(t_d) &= \int_0^\infty f_{go}(t)[1 - F_{stop}(t - t_d)]\, dt, \\ &= \int_0^{t_d} f_{go}(t)\, dt + \int_{t_d}^\infty f_{go}(t)\, [1 - F_{stop}(t - t_d)] \\ &= 1 - \frac{\lambda_{stop}}{\lambda_{stop} + \lambda_{go}}\, \exp[-\lambda_{go} t_d]. \end{aligned}$$

The pdf of the signal-response distribution is given, for $t > t_d$, by

$$\begin{aligned} f_{sr}(t \,|\, t_d) &= f_{go}(t)\, [1 - F_{stop}(t - t_d)]/p_r(t_d) \\ &= \frac{\lambda_{go} \exp[-\lambda_{go} t]\, \exp[-\lambda_{stop}(t - t_d)]}{\left(1 - \dfrac{\lambda_{stop}}{\lambda_{stop} + \lambda_{go}}\, \exp[-\lambda_{go} t_d]\right)} \\ &= \frac{1}{K}(\lambda_{go} + \lambda_{stop})\, \exp[-(\lambda_{go} + \lambda_{stop})(t - t_d)], \end{aligned}$$

with $K = \exp[\lambda_{go}\, t_d](1 + \lambda_{stop}/\lambda_{go}) - \lambda_{stop}/\lambda_{go}$. For $t_d = 0$, we have $K = 1$ and the signal-respond density is identical to an exponential pdf for an independent race between T_{stop} and T_{go}, with parameter $\lambda_{go} + \lambda_{stop}$ and $p_r(t_d) = \lambda_{go}/(\lambda_{go} + \lambda_{stop})$.
For $t \le t_d$, the density simplifies to

$$\begin{aligned} f_{sr}(t \,|\, t_d) &= f_{go}(t)]/p_r(t_d) \\ &= (\lambda_{stop} + \lambda_{go})\, \exp[-\lambda_{go}(t)]. \end{aligned}$$

Computation of the expected value of signal-response RTs yields:

$$\begin{aligned} \mathrm{E}[T_{go} \,|\, T_{go} < T_{stop} + t_d] &= \int_0^\infty t\, f_{sr}(t \,|\, t_d) dt \\ &= \frac{\lambda_{go}\, [1 + (\lambda_{go} + \lambda_{stop}) t_d]}{(\lambda_{go} + \lambda_{stop})\{\exp[\lambda_{go}\, t_d](\lambda_{go} + \lambda_{stop}) - \lambda_{stop}\}}. \end{aligned}$$

In particular, for $t_d = 0$, we obtain $\mathrm{E}[T_{go} \,|\, T_{go} < T_{stop} + 0] = 1/(\lambda_{go} + \lambda_{stop})$, consistent with the density we mentioned above for this value of the stop signal delay.

The exponential distribution does not possess a plausible shape as RT distribution, but it is a special case of the *Weibull* distribution that has just one more parameter. The Weibull is often considered to approximate empirical RT distributions, but the Weibull model has not yet been considered for stop-signal modeling, to our knowledge.

0.4.2 Ex-Gaussian model.

This model, explored by Matzke and colleagues (Matzke et al., 2013), relies on the convolution of an exponential and a normal distribution (ex-Gaussian). The ex-Gaussian distribution is described by three parameters: μ and σ the mean and standard deviation of the Gaussian component, and τ, the mean of the exponential component[10]. It has a positively skewed unimodal shape with μ and σ reflecting the leading edge and τ the tail of the distribution (see Figure 0.7). It often produces an excellent fit to empirical RT distributions.

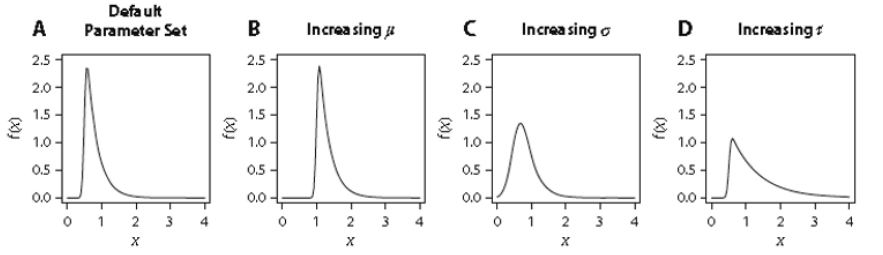

Figure 0.7 Dependence of ex-Gaussian distributional shape on parameter changes. The parameter sets used to generate the distributions are (A) $\mu = 0.5$, $\sigma = 0.05$, $\tau = 0.3$ (default parameter set); (B) $\mu = 1$, $\sigma = 0.05$, $\tau = 0.3$ (increasing μ); (C) $\mu = 0.5$, $\sigma = 0.2$, $\tau = 0.3$ (increasing σ); and (D) $\mu = 0.5$, $\sigma = 0.05$, $\tau = 0.8$ (increasing τ) (from Matzke and Wagenmakers, 2009).

The pdf of the ex-Gaussian is

$$f(t;\mu,\sigma,\tau) = \frac{1}{\tau}\exp\left[\frac{\mu - t}{\tau} + \frac{\sigma^2}{2\tau^2}\right]\Phi\left[\frac{t-\mu}{\sigma} - \frac{\sigma}{\tau}\right] \tag{0.19}$$

where Φ is the standard normal cdf and $\sigma > 0, \tau > 0$. Moreover, as sum of two random variables, the expected value equals $\mu + \tau$ and, by stochastic

[10] Note that here τ is the inverse of the λ parameters in the previous model where the exponential mean was $1/\lambda$.

independence of the component distributions, the variance is $\sigma^2 + \tau^2$. Skewness is determined solely by the exponential component and is equal to $2^{1/3}\tau$ (see also Figure 0.7). The ex-Gaussian model has the theoretical defect of predicting negative RTs with positive probability, but this probability can be made arbitrarily small by shifting the distribution to the right.

The ex-Gaussian stop-signal model assumes separate parameter sets for the T_{go} and T_{stop} distributions, $(\mu_{go}, \sigma_{go}, \tau_{go})$ and $(\mu_{stop}, \sigma_{stop}, \tau_{stop})$. Due to the normal component, no closed-form expressions for $F_{sr}(t)$ and $p_r(t_d)$ are available, but simulation is simple by sampling from the two component distributions and adding the values.

Parameter estimation. While model parameter estimates can be obtained via standard maximum likelihood methods, Matzke and colleagues (Matzke et al., 2013) have also developed a Bayesian estimation method to fit the model to both individual and group data.

First, a uniform prior distribution is assumed for the six parameters of the T_{go} and T_{stop} distributions. These priors are informative in the sense that they cover a wide but realistic range of values informed by results from the stop-signal literature (Band et al., 2003). The prior distributions are then updated by the data to yield the posterior distributions, according to Bayes's rule (without marginal likelihood),

$$\text{posterior} \propto \text{likelihood} \times \text{prior}.$$

For each parameter, the mean, median, or mode of the posterior distribution is taken as a point estimate of the parameter, while the dispersion of the posterior distribution, quantified by the standard deviation or the percentiles, yields information about precision of the parameter estimates: The larger the posterior standard deviation, the greater the uncertainty of the estimated parameter. The posterior distribution for each parameter is approximated via *Gibbs sampling* (Geman and Geman, 1984), a Markov chain Monte Carlo (MCMC) algorithm for obtaining a sequence of observations when direct sampling is difficult (for details, see Matzke et al., 2013). Figure 0.8 illustrates the result for the 3 parameters of the posterior stop-signal pdf.

The Bayesian parametric approach can also handle group data via hierarchical modeling (Gelman and Hill, 2007). Individual parameters are assumed to be drawn from group-level distributions that specify how the individual parameters are distributed in the population. Given that in stop-signal experiments often relatively few observations per participant are available, the hierarchical approach is especially valuable here. For further details about

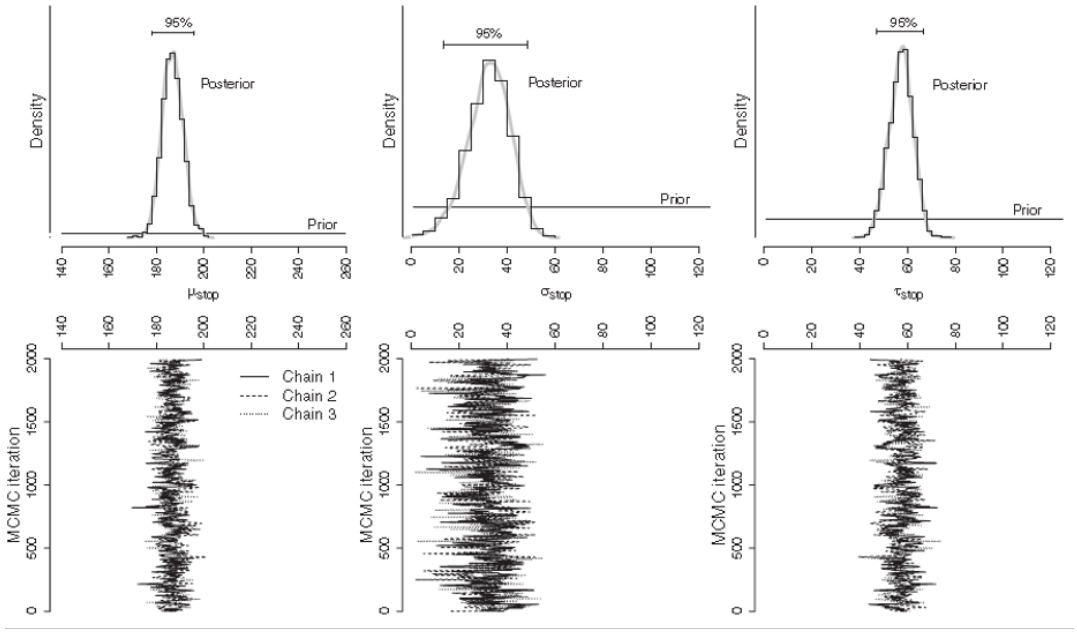

Figure 0.8 The histograms in the top panel show the posterior distribution of the stop signal parameters (synthetic data set). The corresponding thick gray lines indicate the fit of a nonparametric density estimator to the posterior samples. The horizontal black lines at the bottom show the prior distribution of the parameters. The horizontal black lines at the top show the 95% Bayesian confidence interval. The solid, dashed, and dotted lines in the bottom panel represent the different sequences of values (i.e., MCMC chains) sampled from the posterior distribution of the parameters (Gibbs sampling) (from Matzke et al., 2013).

the estimation procedure and accompanying software, we must refer to the original sources (Matzke et al., 2013; Matzke, 2013).

The ex-Gaussian model yields precise information about the unobservable stop signal times but does not attach a specific substantive meaning to the choice of the distribution. In contrast, the following models motivate their distributional form by certain processing assumptions in the stop signal task.

0.4.3 Hanes-Carpenter race model.

The model is based on the linear-approach-to-threshold-with-ergodic-rate (LATER) model purporting to describe the neural mechanism controlling the latency between appearance of a visual target and the start of a saccadic eye movement to the target (Carpenter and Williams, 1995). Introduced in

Hanes and Carpenter (1999), it assumes that the competing go and the stop process rise in a linear fashion to a fixed response threshold. Assuming a fixed response threshold θ, stochastic variability is built into the model by postulating a normally-distributed random rate of rise for going and stopping.

The LATER model assumes a linear rise r of the go process to a fixed threshold, starting from an initial activity level s_0, i.e., $s_0 + r \times t = \theta$ (see Figure 0.9). Assuming r to be the realization of a normally distributed random variable R with mean μ_{go} and variance σ^2_{go}, the above equation leads, after rearrangement, to an expression for the go process random variable T_{go}:

$$T_{go} = (\theta - s_0)/R.$$

Since the distribution of R is given, the pdf of T_{go} follows as (see Colonius et al., 2001)

$$f_{go}(t) = \frac{\theta - s_0}{\sigma_{go}\sqrt{2\pi}\, t^2} \exp\left[-\left(\frac{\theta - s_0}{t} \right)^2 / (2\sigma^2_{go}) \right] \tag{0.20}$$

An analogous pdf is assumed for the stop process T_{stop} with mean μ_{stop} and variance σ^2_{stop} and predictions from the Hanes-Carpenter model are obtained by inserting these distributions into the expression for the signal-respond distribution Equation 0.15 and the analogous expression for $p_r(t_d)$. The model has been tested in several studies. Hanes and Carpenter (1999) reported that the model correctly predicted the probability of successful saccade inhibition as a function of the stop-signal delay as well as the signal-respond distributions. Colonius et al. (2001) found results paralleling those of the non-parametric Logan-Cowan model applied to the same data set, and showed that saccade inhibition is more efficient in response to auditory stop signals than visual stop signals.

Parameter estimation has been performed by minimizing sum-of-squares deviations between observed and predicted data using expressions for the pdfs, by maximum likelihood estimation, and by Monte Carlo simulations (see, e.g., Colonius et al., 2001).

0.4.4 Diffusion race model including its extension to choice RT

In the Hanes-Carpenter model, stochastic variability is implemented across trials by the random rise of activity in going and stopping, but once started, activation accumulates in a linear deterministic fashion within the trial. In contrast, the *diffusion race model* developed in Logan et al. (2014) assumes that both processes are governed by diffusion processes that race against

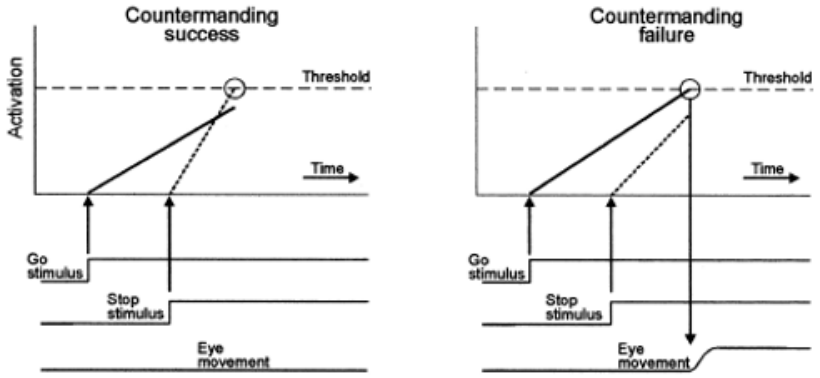

Figure 0.9 The go process (solid line) and the stop process (dotted line) race independently toward their respective thresholds (dashed horizontal line). The thresholds for both processes coincide only for ease of illustration. In stop trials, the stop process is evoked after the go process has begun. Left panel, the go and stop stimuli each trigger a signal rising linearly towards a threshold: if, as here, the stop process rises so fast that it overtakes to go process and reaches threshold first, the saccade is successfully inhibited. Right panel, if the go process reaches threshold first, the saccade fails to be countermanded. (from Hanes and Carpenter, 1999).

each other until the first one reaches a fixed threshold. The concept of a stochastic diffusion (Wiener) process has arguably become the most important component of modeling response times in a wide variety of tasks (e.g. Ratcliff, 1978; Busemeyer and Townsend, 1993; Diederich, 1995; Smith and Ratcliff, 2009; van Zandt et al., 2000) and (for details, see Diederich and Mallahi-Karai, 2018; Smith, 2000).

The diffusion race model assumes a Wiener diffusion process with drift rate ξ, a starting point of zero activation, and a threshold (absorbing boundary) at z. The first-passage time is given by the inverse Gaussian (or, Wald) distribution; for the go process pdf, we get

$$f_{go}(t) = z(2\pi t^3)^{-0.5} \exp\left[-\frac{1}{2t}(\xi t - z)^2\right], \tag{0.21}$$

and for the stop process pdf,

$$f_{stop}(t) = z(2\pi(t - t_d)^3)^{-0.5} \exp\left[-\frac{1}{2(t - t_d)}(\xi(t - t_d) - z)^2\right], \tag{0.22}$$

for $t > t_d$ and zero otherwise.

The model actually tested in Logan et al. (2014) was an extended version of the above, re-introducing across-trial variability by assuming the thresh-

old to be a uniform random variable Z ranging from $z-a$ to $z+a$, with a mean of z and a variance of $a^2/3$. For example, the finishing time of the go process unconditioned over the variable threshold Z results in a go pdf

$$g_{go}(t \mid z, \xi) = (2a)^{-1} \int_{z-a}^{z+a} f_{go}(t \mid z', \xi)\, dz'$$

The context for this model extension was that the authors were interested in modeling a more general paradigm, where participants‘ go response was a decision among stimuli from a set A of possible response alternatives. In this paradigm, participants also produce error RTs (choosing the wrong alternative), and it is well known that the diffusion model with constant threshold cannot predict the often observed “fast error” RT distributions (Smith, 2000).

The diffusion race model is a instantiation of what Logan and colleagues call the *general independent race model*. The latter assumes a double race, first, between a set A of possible go responses and, second, between the winner of the first race and the stop process. Assuming stochastic independence throughout, this implies for the probability that go response i ($i \in A$) will occur given SSD $= t_d$

$$\mathrm{P}\left[\text{response } i \mid t_d\right] = \int_0^\infty f_i(u) \prod_{j \in A, j \neq i} \left(1 - F_j(u)\right) \left(1 - F_{stop}(u - t_d)\right) du,$$

where $F_j(t)$ and $f_j(t)$ ($j \in A$) are the cdf and pdf, respectively, for go response j. The probability that the stop process wins the race is

$$p_{stop}(t_d) = \int_0^\infty f_{stop}(u - t_d) \prod_{i \in A} (1 - F_i(u))\, du;$$

thus, $p_r(t_d) = 1 - p_{stop}(t_d)$.

For the pdf of RTs conditioned on response i, we get the signal-respond distribution

$$f(t|i, t_d) = \left[f_i(t)(1 - F_{stop}(t)) \prod_{j \in A, j \neq i} (1 - F_j(t)) \right] / p_r(t_d).$$

The pdf for T_{go}, the RT to give some response when no stop signal is present, equals

$$f_{go}(t) = \sum_{i \in A} f_i(t) \prod_{j \in A, j \neq i} (1 - F_j(t)).$$

In analogy to Equation (0.9), the signal-respond pdf can be calculated using

$$f_{sr}(t \mid t_d) = \left[\sum_{i \in A} f_i(t) \prod_{j \in A, j \neq i} (1 - F_j(t))(1 - F_{stop}(t - t_d)) \right] / p_r(t_d).$$

This model clearly generalizes the Logan-Cowan race model in covering choice RT paradigms as well. As such, it could be further studied as a non-parametric model, e.g. with an additional assumption of constant SSRT.

However, Logan et al. (2014) were specifically interested in issues of processing capacity. For example, do stop and go process share capacity, or is processing capacity unlimited in the stop signal paradigm? To answer this question, they systematically varied the number of response alternatives and estimated parameters of the race diffusion model. They hypothesized that, under limited capacity, the rate parameter for the stop process should decrease with the number of alternative responses, just as the rate parameters for the go processes do. This is basically what they found using a series of model variants with certain parameters fixed and others free to vary. Moreover, the threshold parameter for the go task increased slightly with the number of alternatives, which is interpreted as subjects adjusting threshold strategically to compensate for the increased noise.

0.5 Parametric dependent race models.

0.5.1 Evidence against independence: the paradox

All models considered up to here were based on assuming both context and stochastic independence. Nevertheless, some recent findings, adding to some earlier ones, have raised serious doubts about the ubiquitous validity of the independence assumptions. A specific independence test is to check that mean signal-respond RTs are monotonically increasing with stop signal delay and that corresponding distribution functions are ordered accordingly (see Fig. 0.10, left panel). In earlier work, we have found some violations of this ordering at short SSDs (e.g., Colonius et al. 2001; Özyurt et al. 2003, see Fig. 0.10, right panel) but evidence remained weak because, typically, observations are sparse at short SSDs. Moreover, Band et al. (2003) investigating the consequences of violations of both context and stochastic independence on stop signal processing estimates via simulation, found severe bias effects on SSRT estimates under some conditions. Recently, in a large-scale survey analyzing 14 experimental studies, Bissett et al. (2021) found serious violations of context independence specifically at short SSDs (i.e. less than 200 ms).

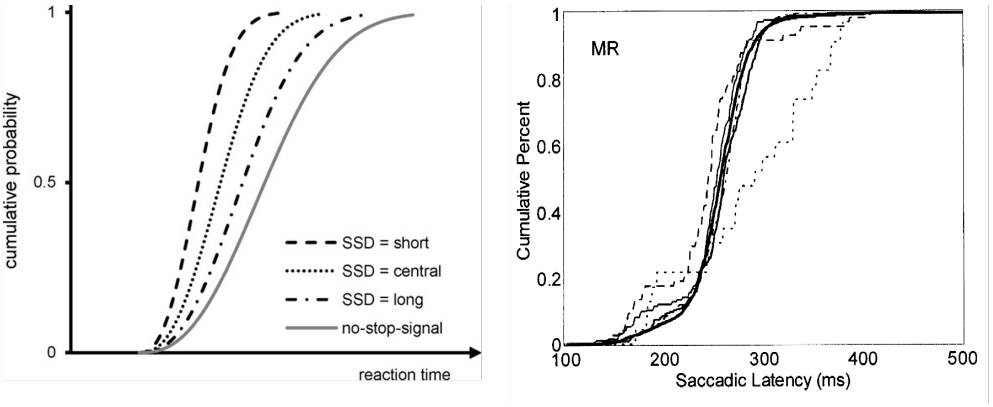

Figure 0.10 Distribution function (cdf) prediction of IND model. Left panel: prediction for signal-respond cdfs ordered by SSD size (from Verbruggen and Logan 2009). Right panel: Observed violation for short SSD = 90 ms (from Colonius et al. 2001). In both panels, the solid line represents the go signal RT distribution.

Such violations are commonly interpreted as refuting context independence, but it seems difficult to tell apart violations of stochastic independence from violations of context independence by experimental tests of behavior. Thus, violations of the former, in addition to or in place of, violations of context independence cannot be ruled out.

Strong evidence against independence comes from seminal findings on the neural underpinnings of response inhibition, and the main impetus for developing race models with dependency arguably comes from these investigations. Studies in the frontal eye fields (FEF) and superior colliculus (SC) of macaque monkeys performing a countermanding task with saccadic eye movements have shown that the neural correlates of go and stop processes produce eye movement behavior through a network of interacting gaze-shifting and gaze-holding neurons (Brown et al., 2008; Hanes et al., 1998; Hanes and Schall, 1995; Paré and Hanes, 2003; Middlebrooks et al., 2020). Specifically, Hanes and colleagues (Hanes and Schall, 1995) showed, first, that macaque monkey behavior in saccade countermanding corresponded to that of human performance in manual stop-signal tasks consistent with the independent model. Then, recording from FEF they isolated neurons involved in gaze shifting and gaze holding that represent a larger circuit of such neurons that extends from cortex through basal ganglia and SC to brainstem (see Figure 0.11).

The question thus arises: How can *interacting* circuits of mutually inhibitory neurons instantiate stop and go processes with (context or stochas-

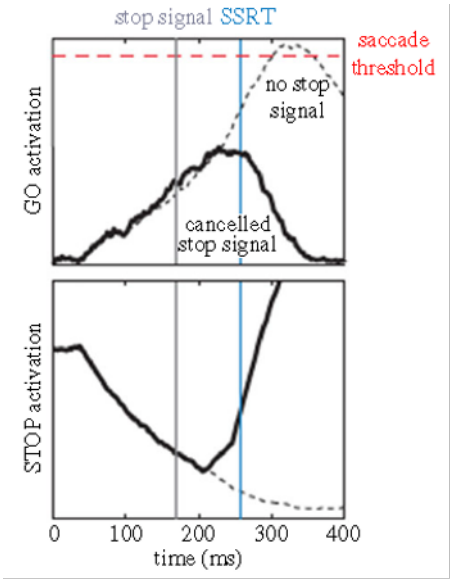

Figure 0.11 Schematic diagram: Activation of the GO unit (upper) and the STOP unit (lower) for trials with no stop signal (dashed lines) and trials with a stop signal that successfully canceled the saccade (solid lines). Saccades are produced when inhibition of the STOP unit is released and the activation of a GO unit reaches a threshold (red dashed line). In response to the stop signal (solid grey line), the STOP unit becomes active, interrupting the accumulation of GO unit activation. This interruption occurs immediately before the stop-signal reaction time (SSRT) (blue line), a measure of STOP process duration derived from the independent race model. From Schall et al. (2017).

tically) independent finishing times? Although it can be argued that behavioral and neural data provide a description on different levels of processing (see below section "linking propositions"), this discrepancy has widely been perceived as a paradox (Boucher et al., 2007; Schall and Godlove, 2012; Schall et al., 2017; Matzke et al., 2018).

In an effort to resolve the paradox, a number of neurally inspired, computationally explicit models have been proposed that will be considered here and in the following section. In Section 0.7, we will present some further behaviorally oriented approaches based on recent concepts of statistical dependence.

0.5.2 Interactive race model

Boucher and colleagues Boucher et al. (2007) developed a relatively simple neural network model, the *interactive race model*, consisting of a go (or move) and a stop (or fixation) unit that accumulate stochastic evidence and race toward a common threshold (arbitrarily set to one). Whichever unit reaches the threshold first determines whether a stop signal trial is signal-inhibit or signal-respond.

The approach, based on a version of the *leaky, competing accumulator model* (Usher and McClelland, 2001; Bogacz et al., 2006), is defined by two stochastic differential equations

$$da_{go}(t) = \frac{dt}{\tau}\left[\mu_{go} - k\, a_{go}(t) - \beta_{stop}\, a_{stop}(t)\right] + \sqrt{\frac{dt}{\tau}}\,\xi_{go}; \tag{0.23}$$

$$da_{stop}(t) = \frac{dt}{\tau}\left[\mu_{stop} - k\, a_{stop}(t) - \beta_{go}\, a_{go}(t)\right] + \sqrt{\frac{dt}{\tau}}\,\xi_{stop}. \tag{0.24}$$

These equations describe the change in activation of the go and stop units, $a_{go}(t)$ and $a_{stop}(t)$, within (an infinitely small) time step dt. Parameters μ_{go} and μ_{stop} denote mean growth rates (*drift rates*) for the go and stop unit, respectively. The leakage parameter, k, prevents the activation from increasing without bound. Interaction between the units is controlled by the inhibition parameters β_{go} and β_{stop} (see Figure 0.12, right panel). The amount of mutual inhibition depends on the instantaneous activation levels, $a_{go}(t)$ and $a_{stop}(t)$, causing a unit with a low activation to have a small inhibitory effect on the other unit. Finally, ξ_{go} and ξ_{stop} are Gaussian noise terms with mean of zero and variance σ^2_{go} and σ^2_{stop}, respectively.

Other parameters in the model capture the non-decision time stages of processing. Stimulus encoding that occurs before go unit and stop unit activation was instantiated is represented as constant delay: D_{go} denotes afferent processing time after the go stimulus is presented, D_{stop} is the latent time after SSD and before the stop unit begins to inhibit the go unit (see Figure 0.12, left panel). Boucher et al. (2007) studied simultaneously recorded behavioral and neural data from two monkeys performing the saccadic stop-signal task (Hanes et al., 1998). Because the above model equations do not possess closed-form solutions, they simulated the model searching for optimal parameter values to minimize deviations of predictions from data. In order to fit neurophysiological data, they first had to decide which parts of the neural populations should correspond to the stop and go units of the interactive model. They noticed:

"The stop-signal task is ideal for investigating the neural control of movement

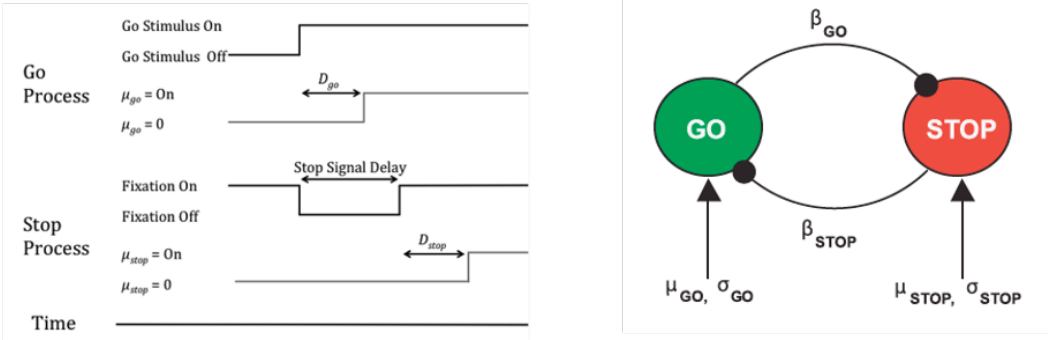

Figure 0.12 Interactive race model. Left panel: timing of events (including afferent delays D_{go} and D_{stop}) (adapted after Logan et al., 2015). Right panel: go and stop unit interaction (with leakage parameter k dropped) (adapted after Boucher et al., 2007).

initiation because it specifies the criteria a neuron must meet to be identified as contributing to controlling saccade initiation. First, the activity in trials when a saccade is made (no-stop-signal or signal-respond trials) must be different from that in trials when no saccade is made (signal-inhibit trials). Second, in stop-signal trials, the activity should begin along the trajectory that would lead to saccade initiation, but on presentation of the stop signal, the activity must be modulated away from that trajectory, and this modulation must occur within the SSRT. Neurons with movement-related and fixation-related activity in frontal eye field and superior colliculus satisfy both of these requirements ."(Boucher et al., 2007, p. 380)[11]

Second, go unit activation was compared with movement neuron activity and stop unit activation was compared with fixation neuron activity. Specifically, for both neurons and model units, activation on signal-inhibit or signal-respond trials was compared with the activity of a subset of latency-matched no-stop-signal trials. No-stop signal trials with response time longer than SSD + SSRT were compared with signal-inhibit trials, because according to the race model, the saccade would have been inhibited had the stop signal been presented. No-stop signal trials with response time shorter than SSD + SSRT were compared with signal-respond trials because, according to the race model, the saccade would have been initiated even if the stop signal had occurred. *Cancel time* was defined as the time at which activation

[11] As the authors point out, determining the quantitative details is a rather subtle task. First, neural activation functions derived from spike trains are converted to spike density functions, as described in Hanes et al. (1998). Although a neural population with a specific function should respond in generally the same way, each neuron may have some idiosyncrasies. Thus, before averaging across neurons, they first had to normalize the spike density function of each neuron by dividing its activity by the peak firing rate in the interval from 20 ms before to 50 ms after saccade initiation on no-stop-signal trials.

on signal-inhibit trials significantly diverged from the activation on no-stop-signal trials relative to SSRT[12] (for further details, we must refer to Boucher et al., 2007, 386 pp.).

In probing the model, Boucher and colleagues first evaluated the ability of the independent race model to account for the observed data. Setting inhibition parameters β_{go} and β_{stop} to zero turns the model into a stochastically independent version, and this resulted in good fits to the behavioral data (inhibition functions, go RT and signal-respond RT distributions). However, since it has no mechanism to shut off the go process so that it does not reach threshold on signal-inhibit trials when the stop process wins, it could not account for the neural data. On the other hand, letting parameters free to vary and utilizing some additional model simulations to estimate go and stop signal cancel times, the authors showed that the interactive model can be fitted simultaneously to both neural and behavioral data. Moreover, by constraining the model parameters in different ways, it turned out that a good model fit depended on two restrictions: (i) activation of the stop unit has to be delayed for a substantial amount of time after the stop signal occurs, i.e., D_{stop} must be rather large (50-70 ms) and (ii) the stop unit must inhibit the go unit much more than vice versa, i.e., β_{stop} has to be much larger (i.e., by an order of magnitude) than β_{go}.

What are the consequences of these findings for the interpretation of SSRT, as measured in the Logan-Cowan independent race model? The parameterized interactive race model implies an additive partition of SSRT: First, stop-signal encoding time, D_{stop}, was between 51 and 67 ms with a small standard deviation (10 to 20 ms). Second, the interval from SSD+D_{stop} until cancel time (interruption of go unit accumulation), called $stop_{interrupt}$, is only about 22 ms, effectively instantaneous. Adding the ballistic interval preceding initiation of the movement (Logan and Cowan, 1984), denoted as $go_{ballistic}$ (about 10 ms), results in the following SSRT decomposition:

$$\text{SSRT} = D_{stop} + stop_{interrupt} + go_{ballistic} \tag{0.25}$$

With SSRT estimates from behavioral data in the range of 80-95 ms, this equation means that most of SSRT is occupied by D_{stop}, during which the go unit is not affected by the stop unit. Boucher et al. (2007) conclude that stopping is a two-stage process consisting of a (relatively long) encoding stage with no interaction and a brief interruption stage during which response preparation is inhibited. This model has been postulated to be a

[12] Cancel time is important in neuroscience because it is an essential criterion for determining whether modulation of neural activity happens early enough to participate in response inhibition (Logan et al., 2015, p. 123).

resolution of the above-mentioned paradox of an independent race at the level of RTs and mutual inhibition at the level of neural activation between gaze-holding and gaze-shifting units (see also Schall et al., 2017).

0.5.3 Linking propositions

The general race model and most of its subclasses do not make a commitment to the underlying computational or neural processes that generate the processing times T_{go} and T_{stop}. The interactive race model, however, has been developed with the aim of connecting go and stop signal processing to the underlying physiology. Given the good understanding of how saccade production is controlled by a circuit of neurons extending from cortex through basal ganglia and superior colliculus to brainstem, the model links the go unit to movement-related neurons and the stop unit to fixation-related neurons in frontal eye fields and superior colliculus (Boucher et al., 2007).

Such *linking propositions* specifying the nature of the mapping between particular cognitive states and neural states have a long history (e.g., Teller, 1984) and they have recently become more popular under the heading of *model-based cognitive neuroscience* (e.g., Forstmann and Wagenmakers, 2015). One motivation for developing linking propositions is the hope to solve the general problem of non-identifiability and model mimicry (see Jones and Dzhafarov, 2014), that exists for behavioral models of choice RT, by identifying the underlying neurophysiology. In the context of the stop signal task, J. Schall and colleagues have thoroughly investigated linking propositions between processing times (T_{go}, T_{stop}) and single-neuron discharges in the frontal eye field, superior colliculus, and ocular motor neurons leading to the interactive race model of Section (0.5.2) and related models (Schall, 2004; Schall and Godlove, 2012; Schall, 2019). Unfortunately, as recently described in Schall (2019), finding a one-to-one mapping between parameters of neural activity and those describing abstract stochastic accumulators (like in race models) seems out of reach at the moment (see also Schall and Paré, 2021).

0.6 Related (non-race) models

0.6.1 Blocked-input model

Starting from the interactive model, Logan and colleagues (Logan et al., 2015) suggested an alternative view on how stopping occurs: the stop unit does not directly inhibit growth of activation of the go unit; rather, the stop

signal activates a top-down process outside the gaze control network that, once reaching a threshold early enough, blocks input to the go unit. Within the dynamics of the interactive model, this means setting go drift rate to zero so that it will not reach its threshold.

The authors defined the units more neutrally as fixation (fix) and movement (move) units because they linked them to gaze-holding and gaze-shifting neurons in a general network, extending from cerebral cortex to the brain stem and being in active balance already at the start of a trial. Modeling steady-state fixation activity implied that eye movements can only occur if activation in the move unit (μ_{move}) and inhibition from the move unit to the fix unit (β_{move}) are large enough to overcome steady-state activation in the fix unit, and if simultaneously inhibition from the fix unit to the move unit (β_{fix}) is not large enough to suppress move activation entirely. For the monkey FEF data from Hanes et al. (1998), these constraints led to equivalent predictions of physiological data for the interactive and the blocked-input model, but the latter model provided a better account of the behavioral data.

By letting certain model parameters vary freely and and keeping others fixed, Logan et al. (2015) compared fits of different versions of the interactive and blocked-input models. Although these models differed strongly with respect to the temporal dynamics of inhibition, they did not show substantial differences in goodness of fit. The authors refer to this result as an instance of "model mimicry" of blocking and inhibiting which can only partially be resolved by considering neurophysiological data.

0.6.2 DINASAUR model

A recent neural network model by Bompas and colleagues (Bompas et al., 2020) tackles the problem of modeling rapid saccadic countermanding from a different background. Their model had originally been developed for the well known phenomenon of *saccadic inhibition*(SI) (Reingold and Stampe, 2002; Bompas and Sumner, 2009; Walker and Benson, 2013).

SI occurs in a paradigm that is, or can be made to be, identical to the stop signal task in all aspects except for the instruction: instead of inhibiting the response upon appearance of a (stop) signal, the participant is instructed to just ignore it and perform the saccade to the target stimulus. The SI effect is manifest as a decrease in the number of saccades observed shortly after (distractor) stimulus onset, compared with baseline conditions (with no signal), with a maximum inhibitory influence occurring around 70–90 ms later (see Figure 0.13).

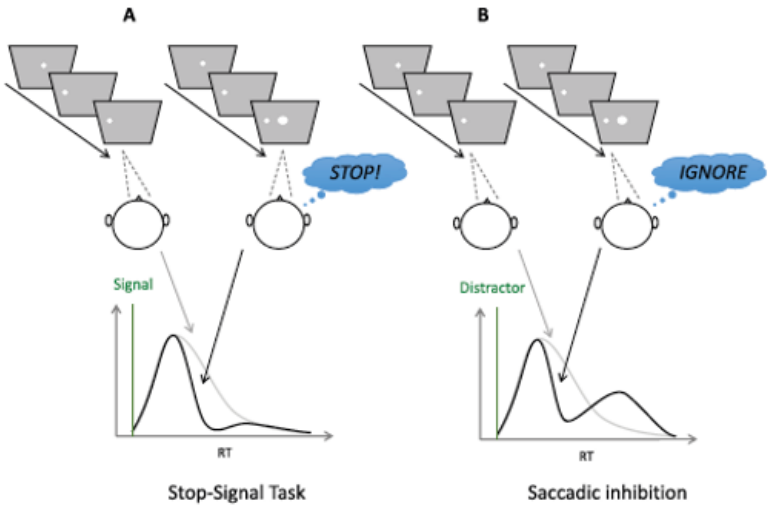

Figure 0.13 Saccadic stop-signal task (panel A) and saccadic inhibition (panel B) paradigms. Both paradigms involve a stimulus jump from center to periphery, sometimes followed by the onset of a central signal (right subpanels above, black lines below), sometimes not (left subpanels, gray lines). The signal onset time is indicated by vertical green lines and the delay between the target jump and the signal is referred to as the stimulus onset asynchrony (SOA). The two tasks differ in the instruction associated with the signal onset: withhold the saccade vs. ignore the signal and perform the saccade. Panel A: Instructions to stop remove slower responses from the RT distribution, but fast responses escape. Panel B: The same visual events associated with an ignore instruction typically produce a dip in the latency distribution, where saccades are delayed and subsequently recover, so that the total number of saccades are about the same between signal present and no-signal distributions (adapted after Bompas et al., 2020).

The authors start from the observation that in the stop signal paradigm, as in SI, fixation and movement neurons receive inputs tightly tied to the visual stimuli (targets and stop-signals) with onsets and offsets leading to step changes some 35 to 50 ms later, preceding inputs from control neurons whose role is to cancel the action plan (Bompas et al., 2020, p. 528). The first part of rapid saccadic countermanding is initially entirely automatic, with slower, top-down endogenous signals built on top of rapid automatic disruption. Their model refers to an approach originally developed by Trappenberg and colleagues (Trappenberg et al., 2001) describing the dynamics of saccadic decision with basic characteristics of exogenous and endogenous

neural signals and lateral inhibition in the intermediate layers of the superior colliculus (SC).

A specific instantiation of Bompas et al.'s model, called 200N-DINASAUR, possesses $n = 200$ nodes representing the horizontal dimension of the visual field, and the average spiking rate A_i of neuron i is a logistic function of its internal state u_i:

$$A_i(t) = 1/(1 + \exp[-\beta\, u_i(t)]).$$

Similar to leaky competing accumulators models, the dynamics of $u_i(t)$ across time depends on normally distributed noise and two types of input received, either external to the map (endogenous or exogenous) or internal via lateral connections (cf. Trappenberg et al., 2001):

$$\tau \frac{du_i}{dt} = -k\, u_i(t) + \frac{1}{n} \sum_j \omega_{ij}\, A_j(t) + I_i^{exo}(t) + I_i^{endo}(t) + N(0, \eta). \quad (0.26)$$

The authors emphasize the distinction between visual events triggering exogenous inputs (i.e. transients tied to visual changes: targets, distractors, or stop signals) not affected by instructions, and endogenous signals (i.e. later, sustained and linked to the instructions) (Bompas et al., 2020, p. 529). Endogenous inputs vary as step functions, while exogenous inputs are transient reaching their maximal amplitude (a_{exo})) at $t = t_{onset} + \delta_{vis}$, and then decreasing exponentially as a function of time, according to the following equation:

$$\tau_{on} \frac{dI_i^{exo}}{dt} = -I_i^{exo}(t) + a_{exo}.$$

Following the Trappenberg et al. model, all inputs have Gaussian spatial profiles (with standard deviation σ). They are maximal at the targeted nodes but also affect nearby nodes. Lateral connections show a Gaussian spatial profile that changes from positive (excitation) at short distance to negative (inhibition) at longer distance according to connection weights ω_{ij} (see Bompas et al., 2020, p. 530).

In the no-signal condition, a single exogenous (visual) transient onset occurs δ_{vis} after target onset, shortly followed by a switch of endogenous support from fixation to target δ_{endo} after target onset. The signal-ignore condition differs from the no-signal condition solely by the presence of a second visual transient, triggered by the signal appearing. When generalizing the model from signal-ignore to signal-stop conditions, only the endogenous input should differ because the visual display is identical and only the instructions differ. As in the blocked-input model, the endogenous input to

the target is switched off (blocked) δ_{endo} after the stop-signal, while the endogenous input to the fixation is switched on again.

Bompas et al. (2020) validate the DINASAUR model in several steps via both simulation and experiment. While the model features a large number of parameters (up to 16), by taking all but 2 of the parameters from the model fit for the SI paradigm in Bompas and Sumner (2009), their simulation was able to reproduce well the typical pattern of results obtained in both paradigms[13](see Figure 0.14).

The model makes two important predictions. Work on SI by Bompas and Sumner (2011) had indicated that dip onset, the time point T_0 where latency distributions diverge, matches the sum of sensory delay δ_{vis} and motor output delay δ_{out} so that $T_0 - SOA$ reflects non-decision time. Moreover, following Boucher et al. (2007), a large portion of SSRT is devoted to non-decision time (the independent processing part, followed by rapid and late inhibition). Thus, Bompas et al. argue that SSRT "...likely behaves like T_0, and therefore we expect the early part of the interference from stop-signals and distractors should be very similar in saccadic inhibition and countermanding." (Bompas et al., 2020, p. 536). Therefore, the first strong prediction of DINASAUR is that the time point at which the RT distribution diverges from the no-signal distribution should be the same under the ignore-signal and the stop-signal instruction (see point T_0 in Figure 0.14, top and bottom right panels).

The second prediction follows from separating exogenous (visual) delay δ_{vis} from endogenous delay δ_{endo}, and from parsimoniously assuming the latter value to be the same in all phases: (a) endogenous support for the target following target onset, (b) the removal of endogenous support for fixation following target onset, (c) the removal of endogenous support for the target following the signal under the stop instruction, and (d) endogenous support returning to fixation following the stop instruction. The prediction then is that extracting these parameters from the no-signal and signal-ignore conditions permits predicting stopping behavior without the need for additional top-down countermanding parameters.

Bompas et al. (2020) found support in three experiments geared toward probing these predictions, but only after adding two amendments to improve fits to the no-signal distribution. The first is to introduce a holding period in order to account for the participants' strategic slowing down in the stop task (*proactive inhibition*). Second, in order to predict "late errors" in the stop signal condition they had to add a parameter for the probability of

[13] Bompas et al. (2020) compare their model with the blocked-input model in much more detail, but we do not go into this here.

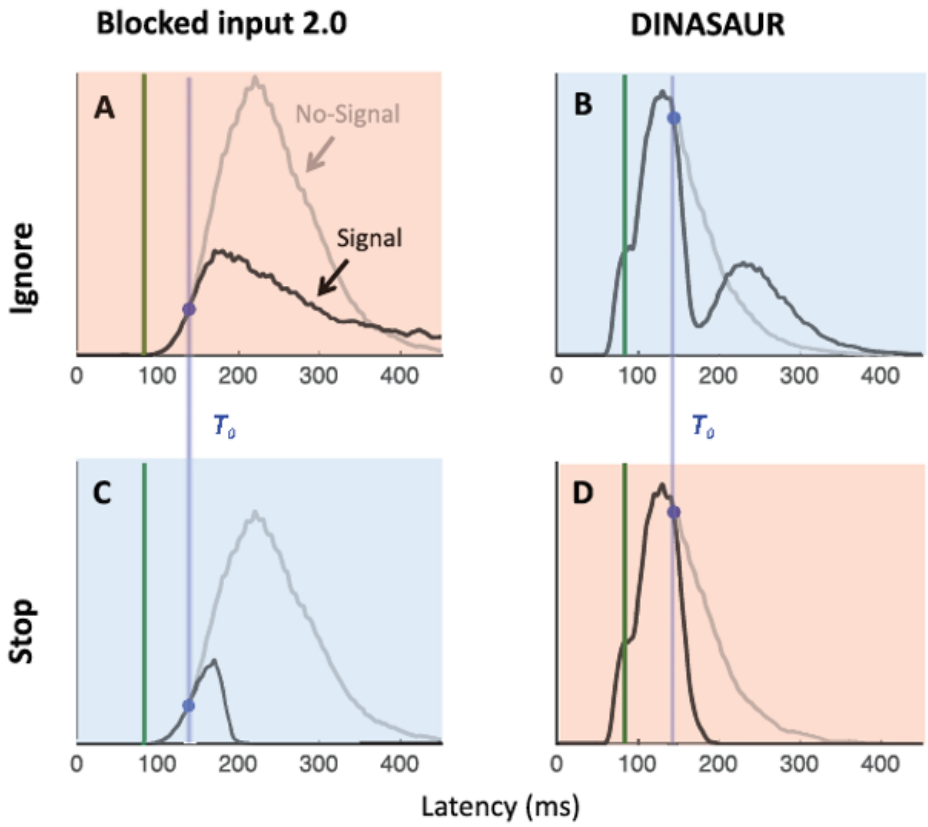

Figure 0.14 Simulated RT distributions for blocked input (left panels) and 200N-DINASAUR (right panels) model for ignore-signal (upper panels) and stop-signal condition (lower panels). The DINASAUR model (with blocked input for stopping) captures well the typical pattern of results obtained in both paradigms. Blocked input 2.0 (with adding automatic fixation activity for ignore conditions) is not able to produce the sharp dips expected from the saccadic inhibition literature. Both models predict a perfect alignment across instructions of the time when the signal RT distribution (black) departs from the no-signal RT distribution (gray), indicated by the blue dots (T_0) and highlighted by blue vertical bars (adapted after Bompas et al., 2020).

not following the stop instruction. This corresponds to the probability of "trigger failures" (see, e.g. Band et al., 2003, and Section 0.8.2 below).

0.6.3 Diffusion-stop model

This model does not implement a race concept either and is related to the blocked-input model closely enough to be mentioned here. In an unpublished paper (Colonius and Diederich, 2001/2021), we address the paradox mentioned at the start of this section by suggesting a diffusion model approach based on Diederich (1997).

The diffusion-stop model assumes a variable growth to a fixed threshold. Rather than claiming separate growths of go and stop signal related activities, it assumes a single diffusion process unfolding over time between two fixed criterion thresholds. The onset of the go signal triggers a growth process represented by a stochastic trajectory drifting towards the upper boundary, threshold (θ_{go}). In the absence of a stop signal, the average trajectory (indicated by the line in Figure 0.15, left panel) has a positive slope resulting in mean saccadic response time determined by the time point corresponding to the crossing of the go threshold. On the other hand, crossing the stop threshold (θ_{stop})) results in a permanent cancellation of the planned movement to the go signal. Figure 0.15 illustrates this mechanism. Presen-

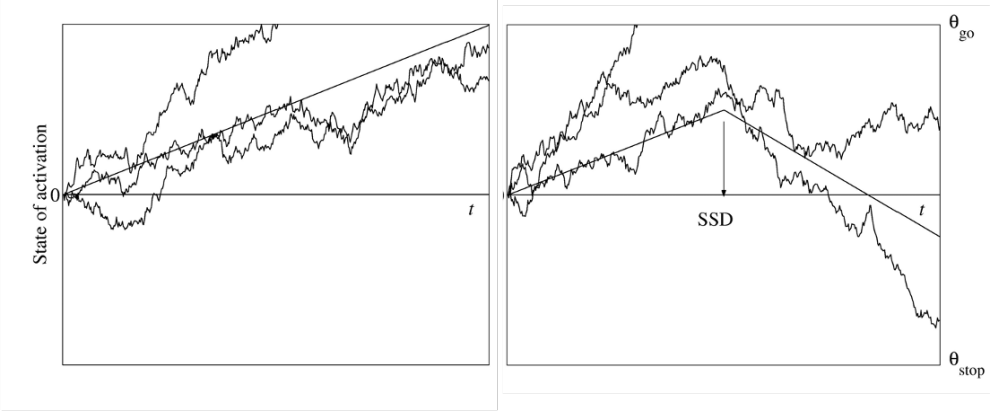

Figure 0.15 Three hypothetical trajectories in the activation space simulated by the diffusion-stop model. Left: in go trials, the drift rate is constant. A saccade is initiated when the trajectory crosses the upper threshold θ_{go} for the first time. The line presents the average trajectory. Right: in stop signal trials, the drift rate switches with presentation of a stop signal at SSD. The saccade is inhibited with certainty once the lower threshold θ_{stop} has been crossed the first time. The average trajectory switches slope from positive to negative at SSD.

tation of a stop signal at point SSD after the go signal shifts the slope of the linear drift to a negative value. Trajectories that have not yet crossed the

upper boundary will then tend in the direction of stop criterion θ_{stop}. Due to stochastic variability, however, individual trajectories may still cross the upper boundary resulting in a response in spite of the stop signal. Note that, like the race model, the diffusion model does not predict different rates of rise in activity for responses in non-canceled trials and in latency-matched no-stop-signal trials. Moreover, consistent with empirical data, the later the stop signal is presented the less likely a successful inhibition of the saccade becomes.

Specifically, the growth process in the diffusion-stop model is represented by a standard Brownian motion (or Wiener) process $A(t)$ with drift rate $\mu(t)$ and two absorbing barriers θ_{go} and θ_{stop}. The process is *time-inhomogeneous* because the drift rate changes with the occurrence of the stop signal at $t =$ SSD:

$$\mu(t) = \begin{cases} \mu_{go}, & \text{if } t \leq \text{SSD}; \\ \mu_{stop}, & \text{if } t > \text{SSD}. \end{cases}$$

The first-passage times are defined as

$$T_{go} = \inf\{t : A(t) \geq \theta_{go} \text{ and } A(\tau) \geq \theta_{stop} \text{ for all } \tau < t\},$$
$$T_{stop} = \inf\{t : A(t) \leq \theta_{stop} \text{ and } A(\tau) \leq \theta_{go} \text{ for all } \tau < t\}$$

with $\theta_{stop} < A(0) < \theta_{go}$.

The model was fit to the data of one subject reported in Colonius et al. (2001) using the finite-state Markov chain approximation of the diffusion process (Diederich, 1997). Observable saccadic reaction time was taken as $SRT = T_{go} + c$ (c a sensorimotor constant) and observable inhibition probability as $\mathrm{P}\,(T_{go} < T_{stop} + \text{SSD})$. Assuming no bias, $A(0)$ was set to zero. Estimated parameters are the drift rate values μ_{go} and μ_{stop}, the distance between go and the stop threshold $\theta_{go} - \theta_{stop}$ and constant c. The model fit (7 data points and 4 parameters) is depicted in Figure 0.16. Note that the diffusion-stop model, in contrast to independent race models, is able to account for the non-monotonic relation between mean response time and stop signal delay.

Measuring the speed of the stop process differs strongly from race models. We define *inhibition time* (IT) as the interval from presentation of the stop signal until the trajectory reaches the stop criterion (lower bound). Thus, IT depends on the momentary level of activity towards the go threshold (represented by the trajectory location) at the time the stop signal is presented. It implies that the average time to cancel a saccade increases with stop signal delay. For example, estimates of IT were 530, 570, and 580 ms for stop signal delays of 70, 100, and 130 ms, respectively. Even if one subtracts some 30

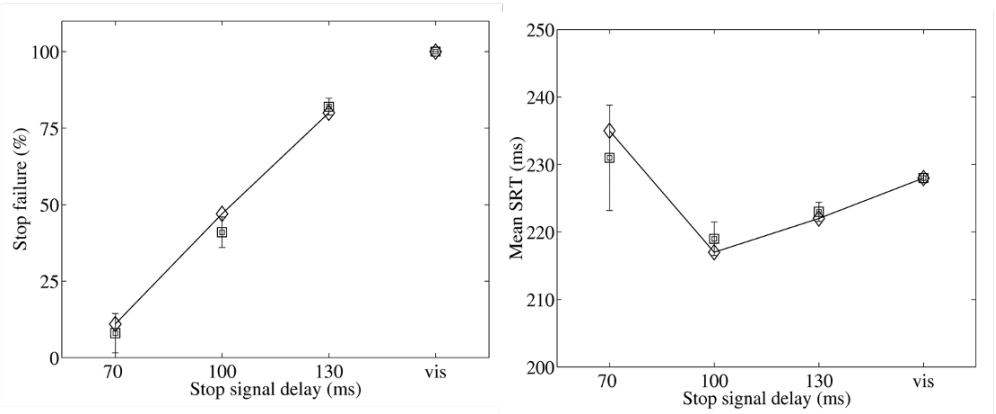

Figure 0.16 Observed data (squares) and predictions (diamonds) from diffusion-stop model for one subject. Left: stop failure probability as a function stop signal delay. Rightmost point refers to no-stop-signal condition. Right: mean saccadic response time as function of stop signal delay.

ms for the latency of the response to the stop signal, the resulting estimates are one half order of magnitude larger than the estimates for SSRT under the race model (100 ms in this case). This discrepancy reflects an important difference between the diffusion-stop and the race model: while both IT and SSRT are initiated by the presentation of a stop signal, termination of IT in the diffusion-stop model indicates that inhibition of the saccade has become certain, whereas termination of SSRT in the race model means that stop signal processing is finished, but actual inhibition of the saccade only occurs if go signal processing has not been terminated earlier. Estimates for IT of about 500 ms in the diffusion-stop model may appear implausible, but it should be noted that this includes the time to suppress the go signal activity completely. If, for example, go and stop signal are presented nearly simultaneously, resulting in a very high probability of successful inhibition, estimates for IT can go down strongly, depending on the relative values of the drift parameters.

Subjects have considerable leeway in performing the countermanding task (proactive inhibition). In the diffusion-stop model, a bias in favoring either stopping performance or response speed is easily accounted for by letting a trajectory in the activation space start from a level closer to the stop criterion or to the go criterion, respectively. Introducing this bias parameter also allows the model to predict sequential effects like a higher probability of canceling a saccade if the movement had failed to be canceled on the previous trial (see Section 0.8).

0.7 Semi-parametric race models

The assumption of context independence is fundamental to the general race model (Section 0.3.1). In addition, stochastic independence has been assumed in all race models discussed so far, with the exception of the interactive race model (Section 0.5.2). Given that this latter model is fully parameterized, one may wonder whether other race models with stochastically dependent "races" can be developed without making strong assumptions about the distributions of T_{go} and T_{stop}.

0.7.1 The role of copulas

It turns out that the concept of *copula* is a natural tool to investigate such dependent race models. Briefly, a copula is a function that specifies how a multivariate distribution is related to its one-dimensional marginal distributions[14]. For stop signal modeling, this means that the bivariate distribution H can be written as[15]

$$H(t,s) = \mathrm{P}\left[T_{stop} \leq t, T_{go} \leq s\right] = \mathrm{C}(F_{stop}(t), F_{go}(s)) \tag{0.27}$$

where C is a bivariate copula that is determined uniquely assuming continuous marginal distributions. Note that a copula specifies the dependency structure without the need to commit to a given distributional family for the marginals, here F_{go} and F_{stop}. For example, letting $u = F_{stop}(t)$ and $v = F_{go}(s)$, copula

$$\mathrm{C}_{IND}(u,v) \equiv u\,v$$

defines stochastically independent race models. Because of the generality of the copula definition, the class of race models based on copulas obviously encompasses all race models with specified marginal distributions.

Example: Farlie-Gumbel-Morgenstern copula. The Farlie-Gumbel-Morgenstern (FGM) copula is defined as

$$\mathrm{C}_{FGM}(u,v) = u\,v[1 + \delta(1-u)\,(1-v)] \tag{0.28}$$

[14] For precise definitions of, and an introduction to, copulas we must refer to the literature Joe (2015); Nelsen (2006); Durante and Sempi (2016); for an introduction in psychological contexts, see Colonius (2016).

[15] Note that in Section 0.7 (only) we write H with the order of marginals switched.

with parameter δ a real-valued constant. It defines a stochastically dependent *semi-parametric* race model with bivariate distribution function

$$\begin{aligned} H_{FGM}(t,s) &= \mathrm{C}_{FGM}(F_{stop}(t), F_{go}(s)) \\ &= F_{stop}(t)F_{go}(s)[1+\delta(1-F_{stop}(t))(1-F_{go}(s))], \end{aligned} \tag{0.29}$$

with parameter δ determining the strength of dependence between T_{go} and T_{stop}. Setting $\delta = 0$ corresponds to the independent race model, negative and positive values of δ to negative or positive dependent models, respectively. It is known that the FGM copula only allows for moderate levels of dependence (e.g. Kendall's tau, $\tau \in [-2/9, 2/9]$)[16].

By inserting specific marginal distributions into a copula, fully parameterized models can be created. For example, with ex-Gaussian marginals with parameters μ and σ for the Gaussian and λ for the exponential component, this results in the ex-Gaussian version of the FGM copula:

$$H_{FGM}(t,s) = F_{stop}(t;\theta_{stop})\, F_{go}(s;\theta_{go})[1+\delta(1-F_{stop}(t;\theta_{stop}))(1-F_{go}(s;\theta_{go}))],$$

where $\theta_{stop} = (\mu_{stop}, \sigma_{stop}, \lambda_{stop})$ and $\theta_{go} = (\mu_{go}, \sigma_{go}, \lambda_{go})$ are parameter vectors, adding up to a total of 7 model parameters including δ.

0.7.2 Equivalence with dependent censoring

Many alternative copula families with relatively simple dependency structures exist and could be investigated. The specific challenge for stop signal race models is, of course, that F_{stop} is unobservable. Fortunately, it turns out that the problem of determining the distribution of non-observable stopping time T_{stop} in the race model is formally equivalent to a problem studied in actuarial science concerned with the *time of failure* of some entity (human, machine, etc.). Recall that *censoring* is a condition in which the failure time is only partially known. For example, *left censoring* occurs if a data point is below a certain value but it is unknown by how much. If the value of the censoring is a random variable, the random censoring time is usually assumed to be statistically independent of the failure time. More recently, however, the determination of failure times under *dependent* random censoring has been considered as well (Wang et al., 2012; Hsieh and Chen, 2020):

Dependent censoring. In medical experiments on tumorigenicity, for example, the failure time of interest, T, is usually the time to tumor onset, which is commonly not observed. Instead only (i) the death (or sacrifice) time of an

[16] FGM copula extensions with a larger dependency range exist but require additional parameters.

animal, serving as the observation time X here, is observed and (ii) whether or not T exceeds the observation time X (at that time, one knows the absence or presence of the tumor). Thus, one can directly estimate the following two functions:

$$G(x) = P(X \leq x) \quad \text{and} \quad p_2(x) = P(X \leq x, T < X), \quad (0 \leq x \leq \infty)$$

by their empirical estimates. With $F(t)$ and $G(x)$ the distribution functions of T and X, respectively, one assumes a copula C

$$\mathrm{C}(F(t), G(x))$$

to specify the dependence between failure time and observation time. Importantly, it has been shown that, under weak assumptions and given the copula, the marginal distribution function F is *uniquely* determined by $G(x)$ and $p_2(x)$ (Wang et al., 2012).

To show the formal equivalence with the dependent race model, we equate distribution $G(x)$ with $F_{go}(s)$ and $F(t)$ with $F_{stop}(t)$. Thus, $p_2(x) = P(X \leq x, T < X)$ corresponds to $P(T_{go} \leq s, T_{stop} + t_d < T_{go})$. Since the latter is not observable, we use the following equality,

$$\begin{aligned}
&P(T_{go} \leq s) - P(T_{go} \leq s, T_{stop} + t_d < T_{go}) \\
&= P(T_{go} \leq s, T_{go} < T_{stop} + t_d) \\
&= P(T_{go} \leq s \,|\, T_{go} < T_{stop} + t_d)\, P(T_{go} < T_{stop} + t_d) \\
&= F_{sr}(s \,|\, t_d)\, [1 - p_r(t_d)],
\end{aligned}$$

showing a one-to-one correspondence between the observable quantities in dependent censoring and the stop signal race model; note that we made use of the correspondence of $p_2(\infty)$ with $P(T_{stop} + t_d < T_{go}) \equiv p_r(t_d)$.

Consequently, the uniqueness result in dependent censoring implies that $F_{stop}(t)$ is uniquely determined in the general race model with a specified copula and that the distribution is amenable to non-parametric estimation methods developed in actuarial science (e.g. Titman, 2014, for a maximum likelihood method). This result is very general and applies to any dependent model, e.g. the FGM model defined in Equation (0.29). Unfortunately, however, a further well-known result from that theory implies that the *numerical value* of the dependence parameter, e.g. δ in the case of the FGM model, is not identifiable in general and, thus, cannot be estimated without specifying the marginals (Titman, 2014; Betensky, 2000). Nevertheless, a sensitivity analysis can be quite revealing about the impact of dependency Wang et al. (2012). In the FGM model, this involves taking a range of dependency parameter values, like $\delta = 0, \pm 0.1, \pm 0.2, \ldots, \pm 0.5$, and probing

how much the predictions generated for the stop signal distribution vary as a function of these values. An application of these results to empirical stop-signal data has not yet been undertaken, however.

We conclude this section with a model featuring extreme stochastic dependency not requiring any numerical parameters.

0.7.3 Perfect negative dependency race model (PND)

In order to resolve the paradox described above, of interacting circuits of mutually inhibitory neurons instantiating stop and go processes in spite of stochastically independent finishing times, we have suggested a race model with negative dependency between go and stop signal processing times (Colonius and Diederich, 2018). It is based on the countermonotonicity copula expressing *perfect negative dependence* (PND) between T_{go} and T_{stop} and is completely parameter-free. The bivariate distribution is defined as

$$H^{-}(s,t) = \max\{F_{go}(s) + F_{stop}(t) - 1, 0\}. \tag{0.30}$$

for all s, t ($s, t \geq 0$). It follows that the marginal distributions of $H^{-}(s,t)$ are the same as before, that is, $F_{go}(s)$ and $F_{stop}(t)$. Moreover, it can be shown that Equation (0.30) implies that

$$F_{stop}(T_{stop}) = 1 - F_{go}(T_{go}) \tag{0.31}$$

holds *almost surely*, that is, with probability 1. Thus, for any F_{go} percentile we immediately obtain the corresponding F_{stop} percentile as complementary probability and vice versa, which expresses perfect negative dependence between T_{go} and T_{stop}[17].

Colonius and Diederich (2018) show that the PND race model is consistent with the empirical data patterns of the stop signal task (Section 0.2) and that one can test the model, at least in principle, against stochastically independent race models. However, experimental studies of the model are not yet available. The PND model arguably constitutes the most direct implementation of the notion of "mutual inhibition" observed in neural data: any increase of inhibitory activity (speed-up of T_{stop}) elicits a corresponding decrease in "go" activity (slow-down of T_{go}) and vice versa.

[17] The relation in Equation (0.31) is also interpretable as "*T_{stop} is (almost surely) a decreasing function of T_{go}*".

0.8 Miscellaneous aspects

0.8.1 Variants of the stop-signal paradigm

Early on, some variants of the standard stop-signal task have been developed in an attempt to gain further insight into response inhibition mechanisms (Logan and Burkell, 1986). Data obtained from these studies are mainly discussed against the background of the independent or the interactive race model. Formal modeling approaches geared toward the specific task variants are rare, however. Here we sketch some results and point out future research goals.

Stop-change paradigm. In stop-change tasks, subjects are instructed to stop the originally planned go response and execute an alternative "change" response (or, "go_2" task) when a signal occurs. A number of experimental and modeling studies suggest that subjects cannot stop and replace a response by simply activating an alternative response. A stop process must inhibit the first go response before the go_2 response can be executed. For some modeling efforts within the multitasking context, we refer to Verbruggen et al. (2008).

Selective-stop paradigm. There are two variants of the selective stop task: in *stimulus-selective* stopping tasks, different signals can be presented and subjects must stop if one of them occurs (valid signal), but not if the others occur (invalid signals); in *motor-selective* stop tasks, subjects must stop some of their responses (e.g. finger press) but not others (e.g. foot press).

For the stimulus-selective task, there are 3 different types of trials: (i) only the go signal is presented, (ii) both the go signal and the stop signal are presented, and (iii) both the go signal and the ignore signal are presented. Mainly, two alternative strategies for stimulus-selective stopping have been discussed within the race model framework: 'Stop then Discriminate' and 'Discriminate then Stop' (Bissett and Logan, 2014). Given that stop and ignore signal are never presented within one and the same trial, it is not obvious that discriminating between stop and ignore signal can naturally be represented as a race. It has been suggested that in the 'Discriminate then Stop' strategy discrimination interferes with go processing, violating the context independence assumption of the independent race model. For this paradigm, further theoretical and experimental work is clearly called for.

Anticipated response inhibition. In anticipated response inhibition (ARI) tasks, participants are required to make a planned response that coincides with a predictably timed event (typically a vertically filling bar) at a predefined stationary target (e.g., horizontal line on the bar). This predefined target requires participants to consistently prepare and initiate movement and is supposed to avoid the 'strategic slowing' often observed in the ordinary stop signal task even when subjects are asked to "respond as soon as possible". Experimental comparisons of ARI tasks with the ordinary stop signal task suggest indeed that SSRT estimates show less bias with this version of response inhibition task (Leunissen et al., 2017). However, a recent study finds violations of context independence due to the nature of the task and suggests a parametric model to take those into account (Matzke et al., 2021)

0.8.2 Modeling trigger failures

All race models assume that go processing (T_{go}) is triggered by presenting the go signal, and stop processing (T_{stop}) by occurrence of the stop signal. However, sometimes no response is registered before the response deadline. These "go omissions" may be due, e.g., to distraction or a lack of attention. For the non-parametric independent race model, a recommendation by Verbruggen et al. (2019), based on extensive simulations, is to assign the maximum observed RT in order to compensate for the lacking responses, when the integration method of estimating SSRT is used.

A more difficult problem arises if the stop signal fails to trigger the stopping process. Simulations have shown that non-parametric estimation methods will overestimate SSRT when trigger failures are present on stop trials (Band et al., 2003). If the probability to respond in stop signal trials (the inhibition function) is larger than zero for small or zero SSDs, this suggests the presence of trigger failures. Unfortunately, there are typically only rather few observations available for very small SSDs making estimates for this probability unreliable. As a pragmatic solution, Verbruggen et al. (2019) suggest researchers include extra stop signals that occur at the same time of the go stimulus but not include these trials in estimating SSRT.

At this time, there is no general solution available for non-parametric race models to estimate the probability of trigger failures in stop signal trials. On the other hand, recent variants of parametric modeling methods provide an estimate of the probability of such trigger failures using a distribution-mixture approach (for details, see Matzke et al., 2019).

0.8.3 Sequential (aftereffects) effects

In a large variety of action control tasks like the stop signal paradigm, participants typically slow down after an error ('post-error slowing'). Several distinct behavioral and physiological explanations have been offered for this observation (Ullsperger et al., 2014), but quantitative models are scarce (though see Dutilh et al. (2012) for a diffusion model approach). One hypothesis attributes slowing to the 'executive system': when it detects an error, it increases control by adjusting the parameters of the perceptual and response system to reduce the likelihood of committing future errors. Consistent with this, subjects often slow down after an unsuccessful stopping in the stop signal task. However, slowing has been observed after successful stopping as well (Verbruggen and Logan, 2008). Bissett and Logan (2012) suggested that the presentation of the stop signal encourages subjects to shift priority from the go task to the stop task, producing longer response latencies after a signal trial and reducing the latency of the stop process. A formal approach has been undertaken by Mohsen Soltanifar and colleagues (Soltanifar et al., 2019). They estimate SSRT separately depending on whether the preceding trial has been a go or a stop trial and then develop a two-state mixture model for the SSRT distribution. They find clear effects of trial type, but further research along these lines is called for.[18] In an earlier development, Angela Yu and colleagues suggested a comprehensive Bayesian inference-based, optimal-control theory for sequential effects (Ma and Yu, 2016) where a control system computes, at any given trial, the probability of a stop signal occurring in the next trial.

0.9 Concluding remarks

In this chapter, we have aimed at characterizing the formal structure of quantitative models for the stop signal task. Possible extensions and generalizations of currently available models have been discussed as well, e.g. the class of semi-parametric race models. Given the rapid increase of experimental studies in this area, presenting the empirical success (or failure) of the various models remains outside the reach of this chapter, however.

A recurring theme concerning model building is the issue of parametric versus non-parametric approaches. On the one hand, the independent, non-parametric race model of Logan and Cowan (1984) (Section 0.3.2), with its straightforward estimation methods for SSRT, has clearly dominated empirical studies up to now, notwithstanding numerous reports of violations

[18] Unfortunately, in this and later papers, these authors always use parametric model versions only.

of some of its assumptions. On the other, the availability of software packages for parameter estimation and model simulation is currently generating a broader usage of parametric race models in applied fields. Increased information about stop signal processing time (beyond the mean), the possibility to more adequately deal with errors in choice paradigms that require discrimination between go signals, and the handling of stop signal trigger failures have been listed among the benefits of the parametric approach (Matzke et al., 2019). It should be mentioned for completeness, though, that it also faces some challenges. There is some arbitrariness involved in the choice of a specific family of distributions for go and stop signal processing times (and, for Bayesian methods, in the choice of priors). For example, the commonly used ex-Gaussian distribution has some features that seem problematic: (i) it has an increasing hazard function, whereas most RT distributions exhibit an increasing and then decreasing (to some constant) hazard function (e.g., Luce, 1986, p. 439), and (ii) it predicts a non-zero probability of realizing negative values. The fact that ex-Gauss distributions often yield good empirical fits does not automatically mean that the ex-Gaussian parameters of the stop signal distribution can be taken as valid description of the inhibitory process. Alternative distribution families have been considered, like the log-normal or the Wald distribution, but detailed studies have sometimes revealed broad parameter identifiability failures for these families (Matzke et al., 2020).

It is difficult to predict what type of behavioral modeling will prevail in the future. In any case, it is obvious that the different variants of the paradigm, like selective stopping, will require going beyond the simple 'race' scheme. Further insight from neurophysiology may suggest more complex mechanisms. A case in point is the two-stage pause-then-cancel (PTC) model by Schmidt and Berke (2017), based on subcortical rodent recordings. As described in Diesburg and Wessel (2021), the first stage is defined by a short-latency "Pause" process that actively delays the go process; it is followed by a slower "Cancel" process, which shuts off ongoing invigoration of the go response. This way, the PTC model tries to disentangle attentional orienting from motor inhibition. The model is clearly at odds with standard, independent race models and calls for an augmented mathematical formalization with more sophisticated quantitative measures for the strength of inhibition.

0.10 Bibliographical notes

While there are a number of early references to the stop signal paradigm (e.g., Lappin and Eriksen, 1966), the first completely developed modeling approach is found in Logan and Cowan (1984). Over the years, a number of review articles have appeared, with different emphases (Band et al., 2003; Verbruggen and Logan, 2009; Logan, 1994; Logan et al., 2014; Matzke et al., 2018; Verbruggen et al., 2019). Platform-independent software to correctly execute the standard stop-signal task by F. Verbruggen is found on GitHub (https://github.com/fredvbrug/STOP-IT). For the anticipated response inhibition task, an open-source program (OSARI) is presented in He et al. (2021).

References

Band, G.P.H., van der Molen, M.W., and Logan, G.D. 2003. Horse-race model simulations of the stop-signal procedure. *Acta Psychologica*, **112**, 105–142.

Betensky, R.A. 2000. On nonidentifiability and noninformative censoring for current status data. *Biometrika*, **81**(1), 218–221.

Bissett, P.G., and Logan, G.D. 2012. Post-stop-signal slowing: Strategies dominate reflexes and implicit learning. *Journal of Experimental Psychology: Human Perception and Performance*, **38**, 746–757.

Bissett, P.G., and Logan, G.D. 2014. Selective stopping? Maybe not. *Journal of Experimental Psychology: General*, **143**(1), 455–472.

Bissett, P.G., Jones, H.M., Poldrack, R.A., and Logan, G.D. 2021. Severe violations of independence in response inhibition tasks. *Science Advances*, **7**, eabf4355.

Bogacz, R., Brown, E., Moehlis, P., Holmes, P., and Cohen, J.D. 2006. The physics of optimal decision making: a formal analysis of models of performance in two-alternative forced-choice tasks. *Psychological Review*, **113**(4), 700–765.

Bompas, A., and Sumner, P. 2009. Temporal dynamics of saccadic distraction. *Journal of Vision*, **9**(17), 1–14.

Bompas, A., and Sumner, P. 2011. Saccadic inhibition reveals the timing of automatic and voluntary signals in the human brain. *The Journal of Neuroscience*, **31**, 12501–12512.

Bompas, A., Campbell, A.E., and Sumner, P. 2020. Cognitive control and automatic interference in mind and brain: a unified model of saccadic inhibition and countermanding. *Psychological Review*, **127**(4), 524–561.

Boucher, L., Palmeri, T.J., Logan, G.D., and Schall, J.D. 2007. Inhibitory control in mind and brain: an interactive race model of countermanding saccades. *Psychological Review*, **114**(2), 376–397.

Brown, J.W., Hanes, D.P., Schall, J.D., and Stuphorn, V. 2008. Relation of frontal eye field activity to saccade initiation during a countermanding task. *Experimental Brain Research*, **190**, 135–151. DOI 10.1007/s00221-008-1455-0.

Busemeyer, J.R., and Townsend, J.T. 1993. Decision field theory: a dynamic-cognitive approach to decision making in an uncertain environment. *Psychological Review*, **100**, 432–459.

Carpenter, R. H.S., and Williams, M.L.L. 1995. Neural computation of log likelihood in control of saccadic eye movements. *Nature*, **377**, 59–62.

Colonius, H. 1990. A note on the stop-signal paradigm, or how to observe the unobservable. *Psychological Review*, **97**(2), 309–312.

Colonius, H. 2016. An invitation to coupling and copulas: with applications to multisensory modeling. *Journal of Mathematical Psychology*, **74**, 2–10. http://dx.doi.org/10.1016/j.jmp.2016.02.004.

Colonius, H., and Diederich, A. 2001/2021. *Measuring the time to cancel a saccade.* unpublished manuscript (available from https://uol.de/en/hans-colonius/my-publications).

Colonius, H., and Diederich, A. 2018. Paradox resolved: stop signal race model with negative dependence. *Psychological Review*, **125**(6), 1051–1058.

Colonius, H., Özyurt, J., and Arndt, P.A. 2001. Countermanding saccades with auditory stop signals: testing the race model. *Vision Research*, **41**(15), 1951–1968.

Diederich, A. 1995. Intersensory facilitation of reaction time: evaluation of counter and diffusion coactivation models. *Journal of Mathematical Psychology*, **39**, 197–215.

Diederich, A. 1997. Dynamic stochastic models for decision making under time constraints. *Journal of Mathematical Psychology*, **41**, 260–274.

Diederich, A., and Mallahi-Karai, K. 2018. Stochastic methods for modeling decision-making. Pages 1–70 of: Batchelder, W. H., Colonius, H., and Dzhafarov, E.N. (eds), *New Handbook of Mathematical Psychology Volume 2*, vol. 2. Cambridge University Press.

Diesburg, D.A., and Wessel, J. R. 2021. The Pause-then-Cancel model of human action-stopping: theoretical considerations and empirical evidence. *Neuroscience & Biobehavioral Reviews*, **129**, 17–34.

Durante, F., and Sempi, C. 2016. *Principles of copula theory.* Boca Raton, FL: CRC Press.

Dutilh, G., Vandekerkhove, J., Forstmann, B.U., Keuleers, E., Brysbaert, M., and Wagenmakers, E.-J. 2012. Testing theories of post-error slowing. *Attention, Perception, & Psychophysics*, **74**, 454–465.

Forstmann, B.U., and Wagenmakers, E.J. 2015. *An introduction to model-based cognitive neuroscience.* 1st edn. NewYork Heidelberg Dordrecht London: Springer-Verlag GmbH.

Gelman, A., and Hill, J. 2007. *Data analysis using regression and multilevel/ hierarchical models.* Cambridge, England: Cambridge University Press.

Geman, S., and Geman, D. 1984. Stochastic Relaxation, Gibbs Distributions, and the Bayesian Restoration of Images. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, **6**(6), 721–741.

Hanes, D.P., and Carpenter, R. H.S. 1999. Countermanding saccades in humans. *Vision Research*, **39**, 2777–2791.

Hanes, D.P., and Schall, J.D. 1995. Countermanding saccades in macaque. *Visual Neuroscience*, **12**, 929–937.

Hanes, D.P., Patterson, W.F., and Schall, J.D. 1998. Role of frontal eye field in countermanding saccades: visual, movement and fixation activity. *Journal of Neurophysiology*, **79**, 817–834.

He, J.L., Hirst, R.J., Puri, R., Coxon, J., Byblow, W., Hinder, M., Skippen, P., Matzke, D., Heathcote, A., Wadsley, C.G., Silk, T., Hyde, C., Parmar, D., Pedapati, E., Gilbert, D.L., Huddleston, D.A., Mostofsky, S., Leunissen, I., MacDonald, H.J., Chowdhury, N.S., Gretton, M., Niktenko, T., Zandbelt, B., Strickland, L., and Puts, N.A.J. 2021. OSARI, an open-source anticipated response inhibition task. *Behavior Research Methods*, https://doi.org/10.3758/s13428–021–01680–9.

Hsieh, J.-J., and Chen, Y.-Y. 2020. Survival function estimation of current status data with dependent censoring. *Statistics and Probability Letters*, **157**(108621).

Joe, H. 2015. *Dependence modeling with copulas*. Monographs on Statistics and Applied Probability, vol. 134. Boca Raton, FL 33487-2742: CRC Press Chapman & Hall.

Jones, M., and Dzhafarov, E.N. 2014. Unfalsifiability and mutual translatability of major modeling schemes for choice reaction time. *Psychological Review*, **121**, 1–32.

Lappin, J. S., and Eriksen, C.W. 1966. Use of a delayed signal to stop a visual reaction-time response. *Journal of Experimental Psychology*, **72**(6), 805–811.

Leunissen, I., Zandbelt, B.B., Potocanac, Z., Swinnen, S.P., and Coxon, J.P. 2017. Reliable estimation of inhibitory efficiency: to anticipate, choose or simply react? *European Journal of Neuroscience*, **45**, 1512–1523.

Logan, G.D. 1994. On the ability to inhibit thought and action: A users' guide to the stop signal paradigm. Pages 189–239 of: Dagenbach, D., and Carr, T.H. (eds), *Inhibitory processes in attention, memory, and language*. San Diego: Academic Press.

Logan, G.D. 2017. Takong control of cognition: an instance perspective on acts of control. *American Psychologist*, **72**(9), 875–884.

Logan, G.D., and Burkell, J. 1986. Dependence and independence in responding to double stimulation: a comparison of stop, change, and dual-task paradigms. *Journal of Experimental Psychology: Human Perception and Performance*, **12**(4), 549–563.

Logan, G.D., and 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., van Zandt, T., Verbruggen, F., and Wagenmakers, E.-J. 2014. On the ability to inhibit thought and action: General and special theories of an act of control. *Psychological Review*, **121**, 66–95.

Logan, G.D., Yamaguchi, M., Schall, J.D., and Palmeri, T.J. 2015. Inhibitory control in mind and brain 2.0: Blocked-input models of saccadic countermandingOn the ability to inhibit thought and action: General and special theories of an act of control. *Psychological Review*, **122**, 115–147.

Luce, R. D. 1986. *Response times: their role in inferring elementary mental organization*. New York, NY: Oxford University Press.

Ma, N., and Yu, A.J. 2016. Inseparability of go and stop in inhibitory control: go stimulus discriminability affects stopping behavior. *Frontiers in Neuroscience*, **10:54**, doi: 10.3389/fnins.2016.00054.

Matzke, D. 2013. Releasing the BEESTS: Bayesian estimation of stop-signal reaction time distributions. *Frontiers in Quantitative Psychology and Measurement*, **4:918**, doi:10.3389/fpsyg.2013.00918.

Matzke, D., and Wagenmakers, E.-J. 2009. Psychological interpretation of the ex-Gaussian and shifted Wald parameters: A diffusion model analysis. *Psychonomic Bulletin & Review*, **16**(5), 798–817.

Matzke, D., Dolan, C.V., Logan, G.D., Brown, S.D., and Wagenmakers, E.-J. 2013. Bayesian parametric estimation of stop-signal reaction time distributions. *Journal of Experimental Psychology: General*, **142**(4), 1047–1073.

Matzke, D., Verbruggen, F., and Logan, G.D. 2018. The stop-signal paradigm. Pages 383–427 of: Wixted, John T. (ed), *Stevens' Handbook of Experimental*

Psychology and Cognitive Neuroscience: Methodology, 4th edn., vol. 4. John Wiley & Sons.

Matzke, D., Curley, S., Gong, C.Q., and Heathcote, A. 2019. Inhibiting responses to difficult choices. *Journal of Experimental Psychology: General*, **148**(1), 124–142.

Matzke, D., Logan, G.D., and Heathcote, A. 2020. A cautionary note on evidence-accumulation models of response inhibition in the stop-signal paradigm. *Computational Brain & Behavior*, **3**, 269–288.

Matzke, D., Strickland, L., Sripada, C., Weigard, A., Puri, R., He, J.L., Hirst, R.J., and Heath, A. 2021, 9 Mar.. Stopping timed actions. *PsyArXiv*, https://doi.org/10.31234/osf.io/9h3v7.

Middlebrooks, P.G., Zandbelt, B.B., Logan, G.D., Palmeri, T.J., and Schall, J.D. 2020. Countermanding perceptual decision-making. *iScience*, **23**(1), https://doi.org/10.1016/j.isci.2019.100777.

Nelsen, R.B. 2006. *An introduction to copulas*. 2nd edn. New York, NY 10013: Springer–Verlag.

Özyurt, J., Colonius, H., and Arndt, P.A. 2003. Countermanding saccades: Evidence against independent processing of go and stop signals. *Perception & Psychophysics*, **65**(3), 420–428.

Paré, M., and Hanes, D.P. 2003. Controlled movement processing: superior colliculus activity associated with countermanded saccades. *Journal of Neuroscience*, **23**, 6480–6489.

Ratcliff, R. 1978. A theory of memory retrieval. *Psychological Review*, **85**, 59–108.

Reingold, E.M., and Stampe, D.M. 2002. Saccadic inhibition in voluntary and reflexive saccades. *Journal of Cognitive Neuroscience*, **14**, 371–388.

Schall, J.D. 2004. On building a bridge between brain and behavior. *Annual Review of Psychology*, **55**, 23–50.

Schall, J.D. 2019. Accumulators, neurons, and response time. *Trends in Neurosciences*, **42**(12), 848–860.

Schall, J.D., and Godlove, D.C. 2012. Current advances and pressing problems in studies of stopping. *Current Opinion in Neurobiology*, **22**, 1012–1021.

Schall, J.D., and Paré, M. 2021. The unknown but knowable relationship between presaccadic accumulation of activity and saccade initiation. *Journal of Computational Neuroscience*, https://doi.org/10.1007/s10827–021–00784–7.

Schall, J.D., Palmeri, T.J., and Logan, G.D. 2017. Models of inhibitory control. *Philosophical Transactions of the Royal Society of London B*, **372**(20160193). http://dx.doi.org/10.1098/rstb.2016.0193.

Schmidt, R., and Berke, J.D. 2017. A Pause-then-Cancel model of stopping: evidence from basal ganglia neurophysiology. *Philosophical Transactions of the Royal Society of London B*, **372**, http://dx.doi.org/10.1098/rstb.2016.0202.

Smith, P.L. 2000. Stochastic dynamic models of response time and accuracy: a foundational primer. *Journal of Mathematical Psychology*, **44**, 408–463.

Smith, P.L., and Ratcliff, R. 2009. An integrated theory of attention and decision making in visual signal detection. *Psychological Review*, **116**, 283–317.

Soltanifar, M., Dupuis, A., Schachar, R. J., and Escobar, M. 2019. A frequentist mixture modeling of stop signal reaction times. *Biostatistics & Epidemiology*, **3**(1), 90–108.

Teller, D.Y. 1984. Linking propositions. *Vision Research*, **24**(10), 1233–1246.

Titman, A.C. 2014. A pool-adjacent-violators type algorithm for non-parametric estimation of current status data with dependent censoring. *Lifetime Data Analysis*, **20**, 444–458.

Trappenberg, T.P., Dorris, M.C., Munoz, D.P., and Klein, R.M. 2001. A model of saccade initiation based on the competitive integration of exogenous and endogenous signals in the superior colliculus. *Journal of Cognitive Neuroscience*, **13**, 256–271.

Ullsperger, M., Danielmeier, C., and Jocham, G. 2014. Neurophysiology of performance monitoring and adaptive behavior. *Physiological Reviews*, **94**, 35–79.

Usher, M., and McClelland, J.L. 2001. The time course of perceptual choice: The leaky, competing accumulator model. *Psychological Review*, **108**, 550–592.

van Zandt, T., Colonius, H., and Proctor, R.W. 2000. A comparison of two response-time models applied to perceptual matching. *Psychonomic Bulletin & Review*, **7**, 208–256.

Verbruggen, F., and Logan, G.D. 2008. Response inhibition in the stop-signal paradigm. *Trends in Cognitive Sciences*, **12**, 418–424. http://doi.org/10.1016/j.tics.2008.07.005.

Verbruggen, F., and Logan, G.D. 2009. Models of response inhibition in the stop-signal and stop-change paradigms. *Neuroscience & Biobehavioral Reviews,*, **33**, 647–661. http://doi.org/10.1016/j.neubiorev.2008.08.014.

Verbruggen, F., Schneider, D.W., and Logan, G.D. 2008. How to stop and change a response: the role of goal activation in multi-tasking. *Journal of Experimental Psychology: Human Perception and Performance*, **34**, 1212–1228.

Verbruggen, F., Aron, A. R., Band, G., Beste, C., Bissett, P G, Brockett, A T, Brown, J W, Chamberlain, S R, Chambers, C D, Colonius, H, Colzato, L S, Corneil, B D, Coxon, J P, Dupuis, A, Eagle, D M, Garavan, H, Greenhouse, I, Heathcote, A, Huster, R J, Jahfari, S, Kenemans, J L, Leunissen, I, Li, C S R, Logan, G D, Matzke, D, Morein-Zamir, S, Murthy, A, Paré, M, Poldrack, R A, Ridderinkhof, K R, Robbins, T W, Roesch, M, Rubia, K, Schachar, R J, Schall, J D, Stock, A-K, Swann, N C, Thakkar, K N, van der Molen, M W, Vermeylen, L, Vink, M, Wessel, J R, Whelan, R, Zandbelt, B B, and Boehler, C N. 2019. A consensus guide to capturing the ability to inhibit actions and impulsive behaviors in the stop-signal task. *eLife*, **8**.

Walker, R., and Benson, V. 2013. Remote distractor effects and saccadic inhibition: Spatial and temporal modulation. *Journal of Vision*, **13**(9), 1–21.

Wang, C., Sun, J., Sun, L., Zhou, J., and Wang, D. 2012. Nonparametric estimation of current status data with dependent censoring. *Lifetime Data Analysis*, **18**, 434–445.

Index