Towards a model-based cognitive neuroscience of stopping – a neuroimaging perspective

Sebastian, A.; Forstmann, B.U.; Matzke, D.

Published in:
Neuroscience and Biobehavioral Reviews

DOI:
10.1016/j.neubiorev.2018.04.011

Citation for published version (APA):

General rights
It is not permitted to download or to forward/distribute the text or part of it without the consent of the author(s) and/or copyright holder(s), other than for strictly personal, individual use, unless the work is under an open content license (like Creative Commons).

Disclaimer/Complaints regulations
If you believe that digital publication of certain material infringes any of your rights or (privacy) interests, please let the Library know, stating your reasons. In case of a legitimate complaint, the Library will make the material inaccessible and/or remove it from the website. Please Ask the Library: https://uba.uva.nl/en/contact, or a letter to: Library of the University of Amsterdam, Secretariat, Singel 425, 1012 WP Amsterdam, The Netherlands. You will be contacted as soon as possible.
Towards a model-based cognitive neuroscience of stopping – a neuroimaging perspective

Alexandra Sebastian\textsuperscript{a,b,*}, Birte U. Forstmann\textsuperscript{b}, Dora Matzke\textsuperscript{c}

\textsuperscript{a} University of Amsterdam, Integrative Model-based Cognitive Neuroscience Research Unit, Amsterdam, The Netherlands
\textsuperscript{b} University Medical Center of the Johannes Gutenberg University Mainz, Emotion Regulation and Impulse Control Group, Focus Program Translational Neuroscience, Department of Psychiatry and Psychotherapy, Germany
\textsuperscript{c} University of Amsterdam, Department of Psychology, The Netherlands

**Article Info**

**Keywords:**
- Decision-making components
- Diffusion decision model
- Independent horse-race model
- Individual differences
- Stop-signal task
- Trigger failures

**Abstract**

Our understanding of the neural correlates of response inhibition has greatly advanced over the last decade. Nevertheless, the specific function of regions within this stopping network remains controversial. The traditional neuroimaging approach cannot capture many processes affecting stopping performance. Despite the shortcomings of the traditional neuroimaging approach and a great progress in mathematical and computational models of stopping, model-based cognitive neuroscience approaches in human neuroimaging studies are largely lacking. To foster model-based approaches to ultimately gain a deeper understanding of the neural signature of stopping, we outline the most prominent models of response inhibition and recent advances in the field. We highlight how a model-based approach in clinical samples has improved our understanding of altered cognitive functions in these disorders. Moreover, we show how linking evidence-accumulation models and neuroimaging data improves the identification of neural pathways involved in the stopping process and helps to delineate these from neural networks of related but distinct functions. In conclusion, adopting a model-based approach is indispensable to identifying the actual neural processes underlying stopping.

1. The neural basis of response inhibition – still controversially debated

How we inhibit inappropriate responses is a central question in cognitive neuroscience. Our understanding of brain networks involved in response inhibition has greatly advanced over the last decade (Aron et al., 2014). Using functional magnetic resonance imaging (fMRI), a fronto-basal ganglia stopping network has been identified comprising right inferior frontal gyrus (IFG), pre-supplemental motor area (pre-SMA), and the basal ganglia including the subthalamic nucleus (Aron, 2011; Bari and Robbins, 2013; Sebastian et al., 2016; Verbruggen and Logan, 2008). Yet, it is still controversially debated which prefrontal region initiates the stopping process and via which pathways it is then implemented.

In particular, the functional role of the right IFG within the stopping network remains highly controversial. It has been linked to outright stopping (Aron et al., 2014; Chikazoe et al., 2009; Dambacher et al., 2014; Lee et al., 2017; Levy and Wagner, 2011; Sebastian et al., 2016), the attentional detection of salient signals (Corbetta et al., 2008; Hampshire et al., 2010; Sharp et al., 2010), updating of action plans (Verbruggen et al., 2010), as well as to multiple cognitive control demands in general (Chatham et al., 2012; Duncan, 2013; Erika-Florence et al., 2014; Hampshire and Sharp, 2015). Similarly, the specific role of the pre-SMA within the stopping network has not yet been clearly identified. While some studies indicate a critical involvement of the pre-SMA in stopping (Floden and Stuss, 2006; Obeso et al., 2017; Sharp et al., 2016; Watanabe et al., 2015), others did not find evidence supporting this view (Dambacher et al., 2014; Verbruggen et al., 2010).

Within the stopping network, a hyper-direct pathway either from the right IFG or the pre-SMA to the subthalamic nucleus bypassing the striatum has been proposed to implement stopping (Aron and Poldrack, 2006; Aron et al., 2007). Using dynamic causal modeling, Rae et al. (2014) recently suggested that the right IFG modulates an excitatory influence of the pre-SMA on the subthalamic nucleus, thereby amplifying the downstream polysynaptic inhibition from the subthalamic nucleus to the motor cortex during outright stopping. Direct evidence from neuroimaging studies for the hyperdirect pathway is, however, scarce, which might also be due to the fact that fMRI of small subcortical basal ganglia nuclei at 3 T is challenging (de Hollander et al., 2017). Moreover, one must be aware that after applying smoothing...
procedures during the preprocessing of functional images, it is quite likely that signals originating from the subthalamic nucleus and the substantia nigra, another subcortical basal ganglia nuclei in close proximity to the subthalamic nucleus, will get mixed (de Hollander et al., 2015). This is important, since studies in Parkinson’s patients indicate distinct functions of these structures in response inhibition. While the subthalamic nucleus was suggested to act as a general brake in a ‘hold your horses’ manner, the substantia nigra is activated by the correct response and inhibits the incorrect response (Frank, 2006; Frank et al., 2007). In line with this view, single-unit electrophysiology in rats suggests that neurons in the subthalamic nucleus respond to stop cues regardless of whether or not the response is successfully inhibited, whereas neurons within the substantia nigra respond only on successful stop trials (Schmidt et al., 2013). Evidence from a human fMRI study suggests that substantia nigra activity during a stop trial predicts the degree of reaction time prolongation on the subsequent go trial which is thought to reflect enhanced cognitive control in a stop-signal task (Boehler et al., 2011b). Despite these findings, the substantia nigra receives very little attention in human fMRI studies. While more widely available 3 T fMRI has repeatedly shown basal ganglia activation particularly in larger nodes such as the striatum, newer technologies such as ultrahigh field fMRI, rapid acquisition techniques, the use of multi-echo sequences, or coil setups tailored towards deep brain structures (de Hollander et al., 2017) as well as the use of probabilistic atlases of subcortical nuclei based on ultrahigh field imaging data (e.g., Kruken et al., 2014) might help in the future to clarify the role of small subcortical nuclei such as the substantia nigra and the subthalamic nucleus in stopping and more generally in cognitive control.

Taken together, while a large body of literature has identified brain regions involved in response inhibition, the specific functional role of the key regions within this network as well as the specific pathways implementing stopping remain to be disentangled. The use of a variety of different methods such as fMRI, functional and effective connectivity measures, transcranial magnetic stimulation, deep brain stimulation, and lesion mapping (e.g., Cai et al., 2012; Dambacher et al., 2014; Floden and Stuss, 2006; Lee et al., 2017; Majid et al., 2012; Mirabella et al., 2012; Rae et al., 2014; Verbruggen et al., 2010; Watanabe et al., 2015; Wessel et al., 2016), complementary neurophysiological studies in monkeys and rodents (e.g., Boucher et al., 2007; Fife et al., 2017; Hanes et al., 1998; Hanes and Schall, 1996; Logan et al., 2015; Schmidt et al., 2013), as well as the development of more complex task designs enabling researchers to more specifically capture response inhibition (e.g., Boehler et al., 2011a; Chikazoe et al., 2009; Majid et al., 2012; Sebastian et al., 2016), have significantly contributed to our current knowledge about the neural underpinnings of response inhibition. One example of how more elaborate task designs have contributed to our understanding of stopping is the usage of selective stopping tasks in neuroimaging studies. While the stop-signal task requires global stopping, modified stop-signal tasks require selectively cancelling certain responses. In functional neuroimaging studies, modified stop-signal tasks have been utilized to disentangle the neural correlates of stopping from brain activity associated with the attentional capture of an infrequent, salient signal (e.g., Boehler et al., 2011a; Cai et al., 2011; Chikazoe et al., 2009; Majid et al., 2012; Sebastian et al., 2016). This revealed differential prefrontal involvement of the right inferior frontal junction and ventrolateral prefrontal cortex in attentional capture and

---

### Fig. 1. Stop-signal task, the underlying horse-race model, and the diffusion decision model.

#### (a) Stop-signal task
- **Go trial:** Go right, Go left
- **Stop trial:** Stop right, Stop left
- Inhibit ✅, Respond ❌

#### (b) Schematics of the horse-race model
- **Process activation:**
  - GO: Inhibit ✅
  - STOP: Respond ❌
- **SSD (Stop Signal Delay):**
  - Go stimulus onsets
  - Stop stimulus onsets

#### (c) Complete horse-race model
- **Signal–respond RT distribution:**
  - Go RT distribution
  - Stop RT distribution
- **SSRT (Stop Signal Reaction Time):**
  - Onset go stimulus
  - Onset stop stimulus

#### (d) Diffusion decision model
- **Response alternatives:**
  - Go right = correct
  - Go left = incorrect
- **Starting point:**
  - Drift rate
- **Boundary separation:**
  - Go right = correct
  - Go left = incorrect

---

A. Sebastian et al.
Neuroscience and Biobehavioral Reviews 90 (2018) 130–136

![Image](https://example.com/image1.png)
response inhibition, respectively, suggesting that right posterior inferior frontal gyrus might be involved in triggering the stop process (e.g., Boehler et al., 2011a; Cai et al., 2011; Chikazoe et al., 2009; Sebastian et al., 2016; see Sharp et al., 2010, for deviating results). Selective stopping tasks are thus one example of how the neural signature of specific subprocesses of stopping performance can be disentangled by combining more elaborate task designs and neuroimaging approaches. Human neuroimaging studies, however, rarely rely on a model-based approach and therefore cannot directly identify the cognitive processes involved in response inhibition. Therefore, we advocate a model-based approach to complement other approaches in studying how stopping is neurally implemented. We propose that a model-based neuroscientific approach is indispensable to identifying the neural pathways implementing response inhibition in humans and disentangling these from other cognitive control functions involved in stop-signal performance.

2. Stop-signal task and models of response inhibition

The stop-signal task is widely used to study response inhibition (Logan and Cowan, 1984). In the stop-signal task, participants perform a two-choice response time (RT) “go” task. On a minority of trials, a stop signal presented shortly after the choice “go” stimulus instructs participants to inhibit their response. Participants typically inhibit their response if the interval between the presentation of the go stimulus and the stop signal (i.e., stop-signal delay; SSD) is short, but cannot inhibit their response if the SSD is long (Fig. 1a).

Stop-signal performance is typically formalized as a race between two independent processes, a go and a stop process. If the stop process finishes first, the response is successfully inhibited; otherwise it is erroneously executed (Logan and Cowan, 1984; Logan et al., 2014) (Fig. 1b). The horse-race model enables the estimation of the latency of the unobservable stop response (i.e., stop-signal response time; SSRT). Standard methods allow for the estimation of summary measures of SSRT as a measure of stopping efficiency (see Matzke et al., 2018, in press, for a recent review of the different estimation methods). Focusing on summary measures, such as means, may, however, result in erroneous conclusions. For example, two groups may have the same mean SSRT, but one group may have an SSRT distribution with a faster leading edge but a larger proportion of slow responses. Ignoring such differences may lead to the incorrect conclusion that the groups do not differ in SSRT.

The horse-race model also allows for the estimation of the entire distribution of SSRTs (Fig. 1c). The underlying horse-race idea remains the same, but both go RT and the SSRT are now treated as random variables (Colonius, 1990; Logan et al., 2014). In particular, the recently developed Bayesian parametric approach (Matzke et al., 2013a; for software, see Matzke et al. 2013b and http://osf.io/5yeh4) enables the hierarchical estimation of SSRT distributions using the assumption that go RTs and SSRTs follow an ex-Gaussian distribution (i.e., convolution of a Gaussian and an exponential random variables). The Bayesian estimation of SSRT distributions is computationally more intensive than standard estimation techniques but it overcomes the shortcomings that are associated with summary measures.

Apart from differences in SSRT latency or variability, individual differences in stop-signal performance might as well reflect differences in the probability of triggering the stop process. Trigger failures can bias the estimation of SSRTs and may cause fictitious group differences. A recent model enables researchers to investigate the relative contribution of trigger failures to stop-signal performance and to correct SSRT estimates for the bias that results from triggering deficiencies (Matzke et al., 2017b). More accurate SSRT estimates will also improve the investigation of the neural correlates of stopping efficiency. Such questions can be addressed by correlating SSRT estimates corrected for trigger failures with brain signal related to successful stopping, preferably while taking into account the uncertainty of the estimates (Ly et al., 2018). Moreover, investigating the neural signature of trigger failures will deepen our understanding of which nodes within the neural stopping network are critically involved in triggering a stop process. Thus, even the simple linkage of behavioral estimates and brain data will most likely benefit from recent advancements in SSRT estimation methods.

While descriptive horse-race models, such as Matzke and colleagues’ (2017b) trigger-failure approach, address whether and when a response is inhibited, they cannot directly identify the cognitive processes involved in response inhibition and do not address how the response is inhibited computationally or neurally (Logan et al., 2015). The race-diffusion model (Logan et al., 2014) is a specific instantiation of the independent horse-race model that aims to identify the cognitive processes involved in stopping by describing how the finishing time distributions of the go and the stop process arise. In particular, the model conceptualizes the go and the stop process as evidence accumulators governed by a diffusion process. The go and the stop process race towards their respective threshold; response inhibition is determined by the first racer that reaches its threshold. The model allows estimation of the time taken up by processes outside the decision-making process, the rate of evidence accumulation, and the threshold for each accumulator. Another prominent model of stopping is the Hanes-Carpenter model (Hanes and Carpenter, 1999) that aims to explain performance in the countermanding paradigm, a saccadic version of the stop-signal task. The model is based on the Linear Approach to Threshold with Ergodic Rate approach (LATER; Carpenter and Williams, 1995), which has been successfully used to identify the processes involved in the initiation of saccades.

The race-diffusion model and the Hanes-Carpenter model enable the identification of the cognitive processes involved in response inhibition but provide no explanation on how stopping might be implemented. This stimulated the development of process models such as the interactive race model and blocked-input models. These models assume that a neuronal stop unit inhibits the growth of activation in the go unit or blocks the input to the go unit, respectively, to prevent it from reaching threshold (Boucher et al., 2007; Cutsuridis, 2017; Logan et al., 2015; Schall et al., 2017).

3. Moving towards a model-based cognitive neuroscience approach

Linking computational models to neurophysiological data provided compelling evidence about neurons instantiating inhibition of eye movements in monkeys (for reviews, see Cutsuridis, 2017; Logan et al., 2015; Schall et al., 2017). Similarly, evidence from rat neurophysiology data provided a deeper understanding of basal ganglia involvement in go and stop processes (for a recent review see Schmidt and Berke, 2017). In human neuroimaging studies, however, a model-based cognitive neuroscience approach (Forstmann et al., 2016, 2011) is still in its early beginnings despite a great progress in mathematical and computational models of response inhibition. Using the stop-signal task in traditional fMRI studies has greatly advanced our understanding of the neural architecture of response inhibition (Aron et al., 2014; Bari and Robbins, 2013; Sebastian et al., 2016), but important issues remain unresolved. The traditional cognitive-neuroscientific approach is incapable of directly capturing most of the processes that affect performance in the stop-signal task and associated neural underpinnings, such as stimulus encoding, rate of information processing, motor preparation and output, speed-accuracy trade-off effects, response bias, and trigger failures (Matzke et al., 2017b; Ratcliff et al., 2016). Accounting for these processes is critical for disentangling neural networks that are specifically involved in implementing response inhibition from those that are involved in other cognitive functions of the decision-making process. This can, however, more efficiently be achieved by applying formal models to behavioral data since such latent cognitive functions and underlying neural processes cannot be captured directly with traditional approaches. Applying computational models involving explicit
stopping mechanisms (e.g., Logan et al., 2015) would provide a means to more directly investigate the neural signature of stopping. The evidence-accumulation framework (e.g., Logan et al., 2014) may also facilitate identification of the decision-making processes directly involved in stopping. Note, however, that sophisticated computational models of response inhibition typically require a large number of stop-signal trials per participant and the simultaneous availability of neurophysiological data. For instance, Logan et al. (2014) used over 2000 stop-signal trials per participant to estimate the parameters of the evidence accumulator corresponding to the stop process, a requirement that cannot be satisfied in neuroimaging studies which rarely rely on more than 200 stop-signal trials.

Thus, methodological challenges related to data quality currently preclude the application of formal models of stopping to human fMRI studies. Therefore, recent behavioral and functional imaging studies in healthy participants and clinical samples have applied models which are not specific to response inhibition, such as standard evidence-accumulation models of choice RTs or the ex-Gaussian decomposition of RT distributions in stop-signal data. Even though these approaches cannot directly tackle the question of the neural signature of stopping, these studies allow valuable insights and conclusions about stopping-related processes. Moreover, such model-based approaches will advance our knowledge about individual differences in stopping performance and will ultimately lead to more comprehensive theories of inhibition deficits in neuropsychiatric disorders.

4. A model-based approach of response inhibition in neurological and psychiatric disorders

The model-based approach is becoming increasingly popular in studies with clinical populations. For instance, by studying diseases in which specific brain structures of interest are affected, researchers can investigate hypotheses about specific pathways within the stopping network or decompose disorder-related deficits. Here we will outline a few examples illustrating the benefits of the model-based approach, using the ex-Gaussian distribution, the descriptive trigger-failure approach (Matzke et al., 2017b) introduced earlier, and the diffusion decision model (DDM; Ratcliff, 1978; Ratcliff et al., 2016). The DDM (Fig. 1d) assumes that dichotomous decisions (e.g., ‘go left’ or ‘go right’) are based on the accumulation of evidence, commencing at a starting point and terminating at a decision threshold that is associated with a particular decision. The inherently noisy diffusion process contributes to erroneous choices and variable RTs. The DDM has four key parameters. First, drift rate as an index of task difficulty or subject ability represents the average amount of evidence accumulated per unit time. Second, boundary separation, i.e., the distance between the decision thresholds, represents response caution and implements the speed-accuracy trade-off. Third, starting point reflects the a priori bias for one of the choice alternatives. Fourth, non-decision time represents the time for peripheral processes such as stimulus encoding and response execution. Consequently, observed RTs are given by the time to diffuse from starting point to boundary, plus the non-decision time (Forstmann et al., 2016; Ratcliff et al., 2016). Evidence accumulation models such as DDM have not only been successfully applied to human data, but also animal data thereby contributing to our knowledge of the internal properties and neurophysiological underpinnings of the decision-making process (e.g., Brunton et al., 2013; Hales et al., 2016; Huk, 2005; Ratcliff et al., 2003; Ratcliff and McKoon, 2008).

Although the DDM and the ex-Gaussian trigger-failure approach require substantially fewer observations than formal models of response inhibition (e.g., Logan et al., 2014, 2015), these models are still demanding of the data. For instance, standard estimation of the trigger-failure approach (Matzke et al., 2017b) requires approximately 200 stop-signal trials per participant. Apart from a fair amount of data, the DDM also requires the availability of a sufficient number of incorrect – error – responses, which requires tailored experimental designs (for a review, see Ratcliff and Childers, 2015). Clearly, this requirement may pose challenges when studying clinical populations, where assessment time is constrained and as a result only a limited number of trials per participant can be obtained. This applies also to fMRI studies in a similar way. This limitation can, however, be addressed by means of Bayesian hierarchical estimation, at least to a certain degree. In contrast to standard parameter estimation, hierarchical estimation acknowledges that participants are not completely independent; this is achieved by explicitly modeling individual differences in parameter values (Gelman and Hill, 2007; Matzke and Wagenmakers, 2009; Shiffrin et al., 2008; Wiecki et al., 2013). By doing so, hierarchical modeling uses information available from the entire group to improve parameter estimation for the individual participants. Hierarchical estimation can therefore result in less variable and on average more accurate parameter estimates and may require fewer trials per participant than standard methods.

4.1. Investigating the involvement of subcortical structures in latent cognitive functions in Parkinson’s patients

Obeso et al. (2014) investigated whether the subthalamic nucleus is involved in adjusting response thresholds and speed-accuracy trade-offs and in context-dependent strategic modulations within a conditional stop-signal task. To this end, they fitted the DDM to go RTs of Parkinson’s patients with left or right subthalatomy, unoperated Parkinson’s patients, and healthy control participants. A classical comparison of RTs and error rates revealed that go RTs of patients with right subthalatomy with the contra-lesional hand were comparable to those of healthy participants, but significantly faster than those of the unoperated patients. Yet, patients with subthalatomy made significantly more discrimination errors than unoperated patients and control participants. Faster RTs and increased error rates indicate that subthalatomy might influence the speed-accuracy trade-off. Taking advantage of the model-based approach, this notion could be corroborated.

Unoperated patients had significantly lower drift rates and higher response thresholds than control participants. Drift rate of patients with right subthalatomy was similar to that of controls and significantly higher than that of unoperated patients, whereas drift rate of patients with left subthalatomy was significantly lower compared to controls. The response threshold of the operated group did not differ from the controls. Patients with right subthalatomy tended to have a lower response threshold than unoperated patients; however, the difference was statistically not significant. In addition, patients with subthalatomy failed to show context-dependent strategic modulation of response thresholds. In contrast to the traditional approach, the DDM analysis enabled Obeso et al. (2014) to directly demonstrate that in particular the right subthalamic nucleus is involved in modulating the rate of information accumulation (i.e., drift rate) and setting the response threshold, thereby influencing the speed-accuracy trade-off on go trials in a stop-signal task.

4.2. Decomposing inhibitory deficits in neurological and psychiatric samples using a model-based approach

To investigate mechanisms of impaired response inhibition in patients with progressive supranuclear palsy and Parkinson’s disease, Zhang et al. (2016) fitted a hierarchical DDM to data from a go/nogo task. In the go/nogo task participants are instructed to react to one stimulus category (go stimulus) but withhold their response in case of a different stimulus category (nogo stimulus). The task is thus similar to the stop-signal task, but the timing of the nogo stimulus and the stop signal differs. While in the stop-signal task the stop signal is presented at some delay after the go signal, the nogo signal in the go/nogo task is presented simultaneously with or instead of the go signal. The go/nogo task and the stop-signal task thus capture action withholding and action
cancellation, respectively, which are related but distinct aspects of response inhibition (Schachar et al., 2007; Sebastian et al., 2013). Accordingly, neuroimaging research has revealed largely overlapping but distinct brain activation patterns for the two tasks (Cieslik et al., 2015; Sebastian et al., 2013; Swick et al., 2011; Zhang et al., 2017). Zhang et al. (2016) demonstrated that while both patient groups displayed impaired response inhibition, these impairments were reflected in different components of the decision making process. Patients with progressive supranuclear palsy displayed a strong response bias towards making a go response and simultaneously had very slow drift rates compared to healthy controls and patients with Parkinson's disease. This might provide a compelling explanation for the apparently paradoxical combination of impaired response inhibition and akinnesia seen in this disease. Patients with Parkinson's disease also displayed a response bias and slow drift rates as compared to healthy controls, however, these impairments were less pronounced than in patients with progressive supranuclear palsy. Non-decision time was shorter in both patient groups as compared to healthy controls, whereas boundary separation was comparable for all groups. The authors suggested that shorter non-decision time might reflect enhanced early sensory processing in the patient groups which might compensate for impaired cognitive function for some extent (Zhang et al., 2016). The model-based approach enabled the researchers to identify disorder-specific differences in latent processes that led to similar behavioral response inhibition impairments. Furthermore, discrimination of patient groups and healthy control participants was significantly improved by using model parameters as compared to common behavioral measures such as error rates and mean RTs, with higher precision and superior signal detection in classification. This study convincingly highlights how formal models can improve our understanding of disease mechanisms and might aid in developing better mechanistic therapies. Zhang et al. (2016) offer explanations how the observed impairments might be linked to abnormal neurophysiology. Direct evidence can, however, only be obtained from model-based neuroimaging studies.

Karalunas and Huang-Pollock (2013) fitted the DDM to go RTs of a stop-signal task of children with and without attention-deficit/hyperactivity disorder (ADHD). While no group differences in boundary separation were present, ADHD patients had slower drift rates reflecting less efficient information processing. Faster non-decision time suggested that ADHD patients may not spend enough time on non-decisional processes such as stimulus encoding and motor execution, which might contribute to poorer overall performance. Finally, an ex-Gaussian decomposition of go RTs revealed that increased RT variability was a result of a higher proportion of RTs in the slow exponential tail of the distribution in ADHD. Mediation analyses revealed that both drift rate and the ex-Gaussian \( \tau \) parameter (representing the higher proportion of slow responses) mediated the relationship between ADHD status and inhibitory control. These results suggest that slower and less efficient information processing in combination with immature stimulus encoding and/or motor preparation and execution contribute to deficits in stop-signal performance in ADHD that are commonly simply attributed to inhibitory deficits. Similarly, schizophrenia has been associated with a larger slow tail of the go RT distribution and an increased rate of trigger failures compared to healthy participants (Matzke et al., 2017a). Moreover, the model-based approach showed that failing to react to the stop-signal and a slower initiation of inhibition both contributed to longer SSRT estimates which would have been overestimated and misinterpreted in the conventional approach. The findings implicate attentional deficits rather than an inhibitory deficit itself in schizophrenia which was further underlined by correlations between the probability of trigger failures and event-related potentials. These studies highlight how a model-based approach enables researchers to better understand the nature of (apparent) inhibitory deficits in neuropsychiatric disorders.

5. First insights from model-based neuroimaging studies

Recently, two neuroimaging studies have adopted a model-based approach utilizing the DDM. To investigate the relationship of individual differences in decision components and neural activation in the stopping network as assessed by fMRI, White et al. (2014) fitted the DDM to go trials of a stop-signal task. Faster processing of the go stimulus (i.e., higher drift rate) correlated with right frontal pole activation on go and successful stop trials. This suggests that frontal pole activation reflects the strength of the go process, since the go process is present on both trial types. Individuals who had faster go processing also activated the stopping network (i.e., right IFG and basal ganglia) more strongly on successful stop trials demonstrating a robust relationship between stimulus processing and response inhibition at the neural level. This relationship is in line with assumptions of the horse-race model: only a fast and efficient stop process can win the race against a fast and accurate go process which requires strong activation of the stopping network. Particularly this relationship might underlie inhibitory deficits in patient populations. Non-decision time negatively correlated with activation of the right parietal cortex during successful stopping. Although this parameter reflects both sensory and motor processes, individual differences are more likely to reflect differences in motor execution hinting at strategic motor slowing in order to improve performance.

By combining DDM-analysis and effective connectivity analyses, Jahfari et al. (2015) identified additional prefrontal-to-visual connections in case of fast and accurate go processing. Higher prefrontal-to-visual connectivity was related to shorter SSRT suggesting that in line with the blocked input model – these connections suppress the fast flow of visual input for the go task, such that the inhibition process can finish before the selection process.

While both studies allow insights and conclusions about neural correlates of stopping-related processes by linking model-based neuroimaging results to computational models, one must note that the DDM was fitted only to go RTs. Therefore, although the evidence resulting from these studies is valuable, it is still indirect. As noted above, applying computational models involving explicit stopping mechanisms (e.g., Logan et al., 2015) would further improve our understanding of the neural basis of stopping. Nevertheless, both studies offer insights in the neural basis of stopping which would not have been obtained using traditional cognitive neuroscience approaches. The findings by White et al. (2014) suggest that a fast and efficient stop process relies particularly on fronto-striatal brain activation. Jahfari et al. (2015) provide first evidence for how stopping might be implemented neurally in humans in line with the blocked input model (Logan et al., 2015). Given these valuable insights, studies like the ones highlighted in this review will hopefully encourage the neuroimaging community to incorporate model-based approaches to advance the field to gain a deeper understanding of the neural signature of stopping and to tackle questions the traditional cognitive neuroscience approach was up to date not able to satisfactorily answer.

6. Limitations

As noted above, methodological challenges related to data quality currently preclude the application of computational models involving explicit stopping mechanisms to human fMRI studies. Another limitation is the time resolution of fMRI on the scale of seconds due to the sluggish nature of the BOLD response. As demonstrated for example using basal-ganglia electrophysiology in rats (Schmidt et al., 2013), the success of response inhibition appears to depend on differences in the millisecond range in the relative timing of the inhibition process on individual trials. This is in line with existing theoretical models (Boucher et al., 2007; Logan et al., 2015). In addition, this is reflected in human scalp-electrophysiology (Wessel and Aron, 2015). BOLD-fMRI may not be sensitive to capture such differences in the millisecond range.
range. This may be one reason why fMRI studies rarely show BOLD differences in key regions such as right IFG and pre-SMA when comparing stop-inhibit to stop-respond trials (e.g., Aron, 2011; Boehler et al., 2010; Congdon et al., 2010). This shortage may be overcome by acquiring simultaneous EEG-fMRI data and thus taking advantage of the fMRI’s high spatial resolution and the EEG’s high temporal resolution (e.g., Schmüser et al., 2014). Note that we can learn most about the cognitive neuroscience of stopping by using a multitude of methods and experimental designs with model-based neuroimaging approaches being one essential part.

7. Concluding remarks

In the present review we highlighted recent advances in model-based analyses of stop-signal performance. We outlined that estimating trigger failures does not only correct SSRT estimates for the bias that results from triggering deficiencies, but also enables researchers to decompose inhibitory deficits in clinical samples. First and very promising model-based neuroimaging approaches lay the foundations for a better understanding of the stopping process. Linking evidence-accumulation models and neuroimaging data improves the identification of neural pathways involved in the process of stopping and delineating these from neural networks of related but distinct functions. The neuroimaging field hitherto lacks the application of computational models involving explicit stopping mechanisms, ideally combined with ultra-high field fMRI, which is thought to improve our understanding of the neural basis of stopping tremendously. Methodological challenges, however, currently precludes this desirable development. In conclusion, we believe that the proposed model-based approach is crucial for the development of a unified theory accounting for individual differences in stopping performance.

Acknowledgments

A.S. was supported by the German Research Foundation (DFG grant SE 2771/1-1). B.U.F. was supported by an ERC grant from the European Research Council and a Vidi grant from the Dutch Organization for Scientific Research. D.M. is supported by a Veni grant (451-15-010) from the Dutch Organization for Scientific Research.

References

Hanes, D.P., Patterson, W.F., Schall, J.D., 1998. Role of frontal eye fields in


A. Sebastian et al.
Neuroscience and Biobehavioral Reviews 90 (2018) 130–136