

Response Suppression by Automatic Retrieval of Stimulus–Stop Association: Evidence from Transcranial Magnetic Stimulation

Yu-Chin Chiu¹, Adam R. Aron¹, and Frederick Verbruggen²

Abstract

Behavioral studies show that subjects respond more slowly to stimuli to which they previously stopped. This response slowing could be explained by “automatic inhibition” (i.e., the reinstatement of motor suppression when a stimulus retrieves a stop association). Here we tested this using TMS. In Experiment 1, participants were trained to go or no-go to stimuli. Then, in a test phase, we compared the corticospinal excitability for go stimuli that were previously associated with stopping (no-go_then_go) with go stimuli that were previously associated with going (go_then_go). Corticospinal excitability was reduced for no-go_then_go compared with go_then_go stimuli at a mere 100 msec poststimulus. Although these results fit with automatic inhibition, there was, surprisingly, no suppression for no-go_then_no-go stimuli, although this should occur. We speculated that automatic

inhibition lies within a continuum between effortful top-down response inhibition and no inhibition at all. When the need for executive control and active response suppression disappears, so does the manifestation of automatic inhibition. Therefore, it should emerge during go/no-go learning and disappear as performance asymptotes. Consistent with this idea, in Experiment 2, we demonstrated reduced corticospinal excitability for no-go versus go trials most prominently in the midphase of training but it wears off as performance asymptotes. We thus provide neurophysiological evidence for an inhibition mechanism that is automatically reinstated when a stimulus retrieves a learned stopping episode, but only in an executive context in which active suppression is required. This demonstrates that automatic and top-down inhibition jointly contribute to goal-directed behavior. ■

INTRODUCTION

Stopping an ongoing or initiated action is an important part of everyday behavior. Most stopping research has focused on outright stopping, which is triggered by an external signal and is regulated by an “executive control” network including the right inferior frontal gyrus (rIFG), the pre-SMA, the basal ganglia, and the primary motor cortex (for reviews, see Aron, 2011; Boehler, Appelbaum, Krebs, Hopf, & Woldorff, 2010; Chambers, Garavan, & Bellgrove, 2009; Verbruggen & Logan, 2008b). However, recent evidence suggests that stopping can also be triggered in an unintentional (bottom-up) fashion by the retrieval of previously acquired stimulus–stop associations, leading to the slowing down of responses—so called “automatic inhibition” (Verbruggen & Logan, 2008a). This phenomenon could be practically useful. For example, it has been shown that when pictures of alcoholic drinks were paired with stopping, there was a subsequent reduction in alcohol-related consumption (Houben, Nederkoorn, Wiers, & Jansen, 2011). It is therefore important to better understand the neurophysiological mechanisms underlying automatic inhibition. Here we do so by using TMS to probe the corticospinal excitability of response representations during task performance.

The original study by Verbruggen and Logan (2008a) hypothesized that automatic inhibition develops when stimuli are consistently paired with stopping during a training phase. To measure the effects of automatic inhibition, they compared the RTs for stimuli that were no-go in the training phase but became go stimuli in a test phase (i.e., no-go_then_go, the “inconsistent condition”) with go stimuli that stayed go (i.e., go_then_go, the “consistent condition”). In a series of experiments, they found that responding to go stimuli in the test phase was slower for the inconsistent items compared with the consistent items. On the basis of these findings, Verbruggen and Logan (2008a) proposed that response inhibition could be triggered automatically by the retrieval of stimulus–stop associations after practice.

A follow-up neuroimaging study with a similar paradigm demonstrated that inconsistent go items (that were associated with stopping in the training phase) activated the rIFG (Lenartowicz, Verbruggen, Logan, & Poldrack, 2011). As the rIFG is a key component of neural circuitry for inhibitory control (e.g., Aron, Behrens, Smith, Frank, & Poldrack, 2007; Chikazoe, Konishi, Asari, Jimura, & Miyashita, 2007), the finding suggests that the retrieval of stimulus–stop associations (i.e., automatic inhibition) has its counterpart in a reinstatement of inhibitory control at the neural level. However, because activation of the rIFG

¹University of California—San Diego, ²University of Exeter

has been associated with other functions, such as attentional orienting, reversal learning, conflict, and working memory (e.g., Dodds, Morein-Zamir, & Robbins, 2011; Levy & Wagner, 2011; Hampshire, Chamberlain, Monti, Duncan, & Owen, 2010), the activation of rIFG in this paradigm cannot be taken as proof that the retrieval of stimulus–stop associations necessarily comes along with inhibitory control over the motor system.

Here, we tested whether this response slowing was indeed because of automatic inhibition in the motor system triggered by the reversal of the stimulus go/no-go mapping. We used the same go/no-go paradigm as Verbruggen and Logan (2008a, Experiment 1) along with TMS to probe corticospinal excitability. On each trial in the training phase, participants made a semantic judgment about a word (e.g., living vs. nonliving). The category determined whether the participant should go or stop (e.g., living = go, and nonliving = no-go). In two groups of participants, we manipulated the consistency of the stimulus go/no-go mapping across a training phase and a test phase. For the consistent group, the go/no-go mapping in the test phase corresponded to the go/no-go mapping of the training phase. For the inconsistent group, the mapping was reversed; that is, the stimuli that were associated with stopping (no-go) in the training phase required a go response in the test phase, and vice versa. In the test phase, we applied TMS to the contralateral motor cortex of the responding hand (only one hand was used to respond on go trials) and measured the motor-evoked potentials (MEPs) at precisely 100, 150, 200, and 250 msec after the SOA. The magnetic pulse leads to action potentials in pyramidal cells in the primary motor cortex, and these cells project to the spinal cord and ultimately to the muscles of (in this case) the hand. The MEP—a measure of corticospinal excitability—reflects the net influence of cortical and subcortical motor regions on the specific response channel that is measured at the hand with EMG.

On the basis of the previous behavioral results (Verbruggen & Logan, 2008a), the automatic inhibition hypothesis makes two clear predictions for the MEP pattern. First, there will be reduced corticomotor excitability soon after the presentation of stimuli that were previously associated with stopping. Specifically, we expected (a) in the consistent condition, lower MEPs for no-go (i.e., no-go_then_no-go) trials than go (i.e., go_then_go) trials and (b) in the inconsistent condition, lower MEPs for go (i.e., no-go_then_go) trials than no-go (i.e., go_then_no-go) trials. In other words, we predicted an interaction between trial type and mapping. Furthermore, we hypothesized that the MEP reduction, because of “automatic inhibition,” would occur at the early prereshponse initiation SOAs of 100 or 150 msec (cf. Jahfari, Stinear, Claffey, Verbruggen, & Aron, 2010). This prediction was motivated by a recent EEG study that observed a reduction in the stimulus-evoked potential very early in time (61–104 msec poststimulus) for no-go stimuli at the end of a go/no-go training, which is possibly mediated by a

fast-acting inhibitory mechanism (Manuel, Grivel, Bernasconi, Murray, & Spierer, 2010).

Second, at the later SOAs of 200 or 250 msec, we predicted a main effect of mapping and an interaction between trial type and mapping. That is, MEPs will be larger for go trials than for no-go trials, but this difference should be more pronounced in the consistent group than in the inconsistent group. In the consistent group, the MEPs for go items should start to increase at these later SOAs as motor facilitation starts to build up (cf. Jahfari et al., 2010). However, in the inconsistent condition, this buildup will be delayed because of the time required to recover from putative automatically triggered inhibition, which could account for the slower RTs.

Alternatives to the “automatic inhibition” account are mentioned in the General Discussion. However, to anticipate the outcome of this experiment, the findings were consistent with a modified version of the “automatic inhibition” hypothesis, according to which stimulus–stop associations do suppress motor output, but only in settings in which active suppression is occasionally required.

EXPERIMENT 1

Methods

Participants

Twenty-six volunteers (18–23 years old, mean age = 19 years, 11 women) from the University of California, San Diego, participated in this study for monetary compensation (\$15/hr). All participants were native English speakers, had normal or corrected-to-normal vision, passed a TMS safety screening, and reported no history of neurological impairment. Written informed consent was attained according to an institutional review board protocol. Fourteen volunteers participated in the consistent-mapping condition and 12 in the inconsistent-mapping condition. Four participants in the consistent condition and one participant in the inconsistent condition were excluded due to having outlier data (i.e., greater than two standard deviations above the condition mean).

Apparatus and Stimuli

The experiment was run on an Apple iMac with a 17-in. monitor and a numeric keypad. The keypad was positioned vertically to allow for a lateral extension of the right index finger, which was used for responding in the task (cf. Claffey, Sheldon, Stinear, Verbruggen, & Aron, 2010). This movement is optimal for attaining good EMG signals from the first dorsal interosseous (FDI) muscle. Stimulus presentation and behavioral data collection were controlled using custom MATLAB code (The MathWorks, Natick, MA) and the Psychophysics Toolbox 3 (Brainard, 1997). A list of 144 words drawn from Arrington and Logan (2004) was used in this experiment. We used two tasks (see below). Seventy-two words were used in a living/nonliving

Table 1. Word Length and Word Frequency (per million; see Kučera & Francis, 1967) for the Different Stimulus Categories

Stimuli	Word Length		Word Frequency	
	Living	Nonliving	Living	Nonliving
<i>Small</i>				
<i>M</i>	5.3	5.9	10.6	11.3
Range	3–9	3–11	2–37	1–34
<i>Large</i>				
<i>M</i>	5.0	5.6	10.9	11.2
Range	3–8	3–9	1–32	2–32

categorization task (36 words in each category), and 72 were used in a smaller/larger categorization task (36 words in each category). Word length and word frequency data for the stimuli are shown in Table 1. Each word was presented 12 times during the training phase and four times during the testing phase.

Procedure

Participants were assigned to one of the stimulus–go/no-go mapping conditions (i.e., consistent or inconsistent). Half of the participants in each mapping condition performed the living/nonliving task, and the other half performed the small/large task. See Figure 1A for the experimental design. The category–go/no-go mapping (e.g., living = go, and nonliving = no-go) was pseudo-randomized across participants.

After providing informed consent, participants were seated in front of the computer. First, the TMS motor thresholding procedure was performed (see below). This was followed by the behavioral training phase. In this phase, the go/no-go-mappings were explained verbally, after which the participants performed 12 blocks of 72 trials

without TMS. Each trial during training began with a central fixation cross for 300 msec, followed by a blank screen for 500 msec, followed by the stimulus word for 1000 msec. On go trials, participants had to respond before the word disappeared (1000 msec) by pressing the keypad with their right index finger. On no-go trials, they had to refrain from responding. If they did not refrain (no-go trials) or failed to respond in time (go trials), a beep was played to signal the error. Participants were instructed to respond as fast as possible without making more than 5% of errors. The next trial started after an intertrial interval of 400 msec. Participants received a short break after every two blocks during the training phase.

Immediately after the training phase, the test phase started. Participants in the inconsistent group received new instructions that explained the new go/no-go mapping. Participants in the consistent group were told that the go/no-go mapping remained the same. All participants were informed that TMS would be delivered during the test phase. There were four blocks of 72 trials, with a short break given between each block. The trial progression was the same as the training phase, except that the intertrial interval was now 2200 msec to allow time for the TMS machine to recharge. During each trial, one TMS pulse was delivered at one of the time points (i.e., –200 [baseline], 100, 150, 200, and 250 msec) after the onset of the stimulus word. In each block, TMS was delivered on 8 trials at –200 msec (i.e., before the word onset) and 16 trials at each of the other time points. The ordering of different TMS time points was randomized for each block and each participant. In other words, the occurrence of a pulse did not predict whether subjects should go or no-go.

TMS Procedure

Surface EMG was recorded from the FDI muscle of the right hand using pairs of 10-mm silver electrodes. A ground electrode was placed at the wrist of the right

Figure 1. (A) Experimental design. In the consistent group, the response mapping rules (e.g., living = go, non-living = no-go) remained the same in the training and the test phase whereas in the inconsistent group the response mapping rules of the training phase were reversed in the test phase (e.g., living = go in training, but no-go in test). The red/green color is for illustration only. (B) In the test phase, participants received TMS over their left primary motor cortex while EMG was recorded from the muscle of their right (responding) hand. The gray rectangle indicates the placement of the keypad.

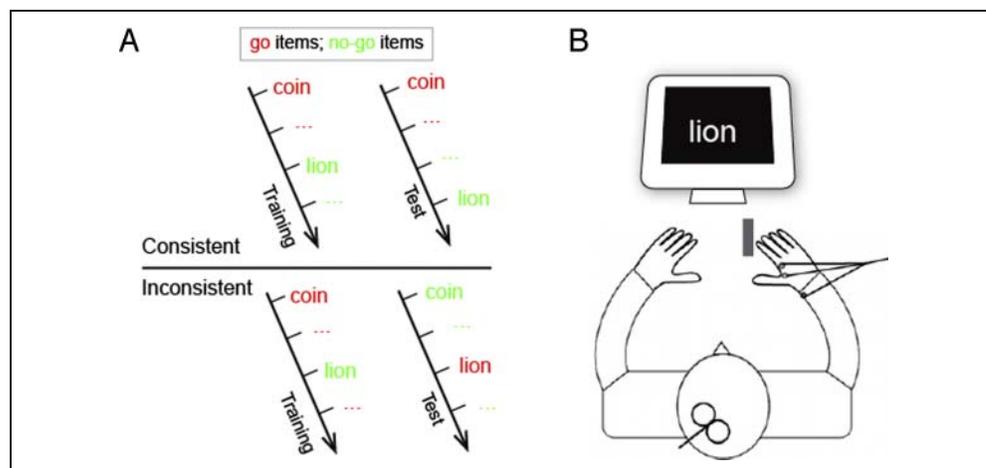
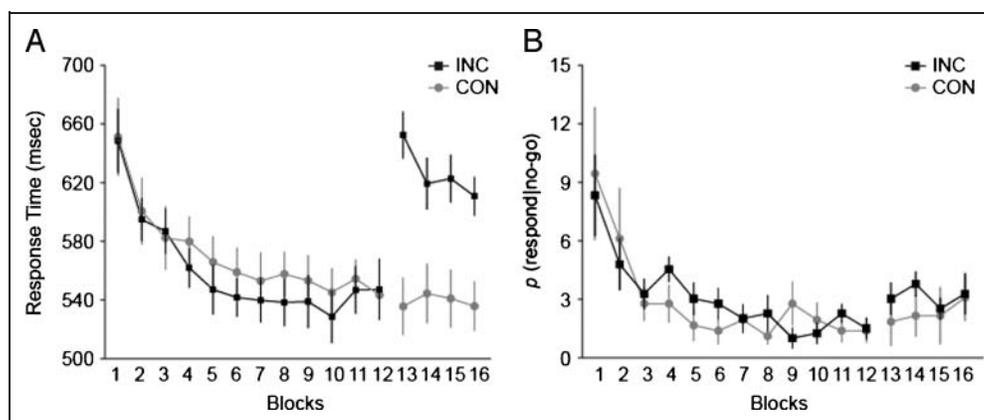


Figure 2. Behavioral performance in the training phase (Blocks 1–12) and test phase (Blocks 13–16). (A) RTs for correct go trials. (B) Probability of responding during no-go trials or p (respond|no-go). Both behavioral measures indicated sufficient learning of stimulus go/no-go mapping during the training phase for both groups of participants. In the test phase, there were prolonged RTs for the inconsistent group but not for the consistent group.



hand (see Figure 1B). The EMG signal was amplified, filtered with a 30-Hz to 1-kHz band-pass filter and a 60-Hz notch filter (Grass QP511 Quad AC Amplifier System, Grass Technologies, West Warwick, RI) and digitized at a rate of 2 kHz (CED Micro 1401 mk II acquisition system).

Recording of the EMG sweep started simultaneously with fixation onset and continued for 2 sec. A Magstim 200² system (Magstim Company, Whitland, Dyfed, UK) was used to deliver TMS pulses via a figure-eight coil. To find the location (“hot spot”) in the left primary cortex for eliciting the best MEPs in the right FDI muscle in the left primary cortex, the coil was initially placed at a point 5 cm lateral and 2 cm anterior to the vertex approximately 45° to the midsagittal line. The coil was incrementally repositioned while administering a single TMS pulse to locate the position that produced reliable MEPs in the right FDI when participants were at rest. The hot spot was marked on a swim cap worn by the participant to ensure consistent coil placement throughout the test phase. The direction of the induced current in the coil was posterior to anterior.

After the hot spot was located, the resting motor threshold for the FDI muscle was determined to the nearest 1% of stimulator output and defined as the lowest stimulus intensity required to elicit MEPs with peak-to-peak amplitude greater than 50 μ V in at least 5 of 10 trials (Rossini et al., 1994). During the experiment, 120% of the resting motor threshold was used as the stimulus intensity.

For each TMS trial, the peak-to-peak amplitude of the MEP was calculated and was used in the statistical analysis. Trials were rejected if the maximum EMG trace during the pretrial epoch (–100 to 0 since TMS onset) exceeded 0.1 mV or the MEP was less than 0.2 mV (mean number of trials rejected = 3.6). The MEP amplitude in each condition was trimmed by removing the top and bottom 10% of the data to reduce the effect of outliers on the mean calculation. To normalize data across participants, the MEP in each condition was divided by each participant’s baseline MEP. The raw baseline MEPs for the consistent and inconsistent mappings were 0.77 ($SE = 0.06$) and 0.63 ($SE = 0.09$), respectively; this difference was not reliable, $t(19) = 1.24$, $p > .12$. Finally, to verify that the FDI muscle was

equally at rest before TMS in each condition, the root mean square of the EMG trace from 200 to 100 msec before TMS onset was calculated.

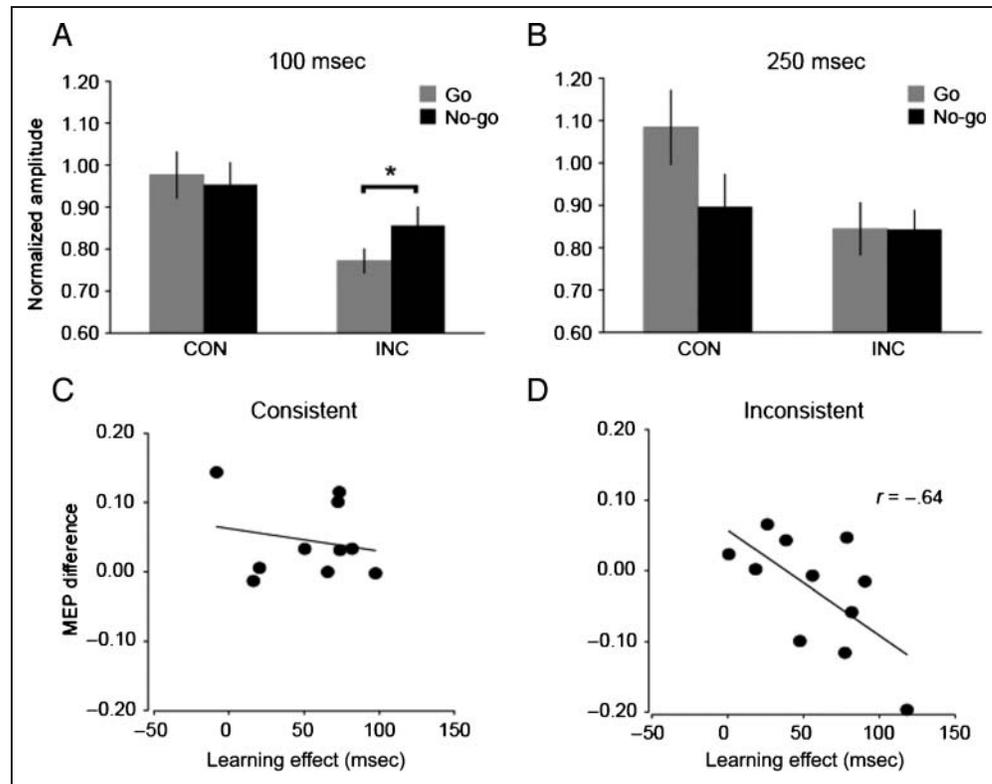
Results

Behavioral Performance

During the training phase, participants performed the go/no-go task without TMS. Over the course of 12 blocks of training (~30 min), RTs on go trials (RT) and probabilities of responding on no-go trials [$p(\text{respond}|\text{no-go})$] decreased for both groups. This suggests that participants learned the stimulus–go and stimulus–stop associations (see Figure 2). This conclusion was supported by a significant main effect of Training Block, $F(11, 209) = 30.8$ for RT; $F(11, 209) = 7.8$ for $p(\text{respond}|\text{no-go})$ (both $ps < .0001$). As expected, the main effect of mapping condition and the interaction between Mapping Condition and Training Block did not reach significance for RT and for $p(\text{respond}|\text{no-go})$ (all $ps > .7$) in the training phase. This confirms that both groups had similar behavior in the training phase.

In the test phase, there was a significant main effect of Mapping Condition: RTs were slower for the inconsistent condition in which the go/no-go mapping was reversed (mean RT = 625 msec) than in the consistent condition (mean RT = 538 msec), $F(1, 19) = 12.9$, $p < .01$. There was no main effect of Block for RTs, $F(3, 57) = 2.02$, $p > .1$, but there was a significant interaction between Block and Mapping Condition, $F(3, 57) = 2.82$, $p < .05$. Figure 2A shows that the interaction effect was driven by the significant effect of block for the inconsistent mapping condition, $F(3, 30) = 3.84$, $p < .02$, but not for the consistent condition ($p > .8$). This suggests that the performance already reached asymptote at the beginning of the test phase in the consistent condition. By contrast, in the inconsistent condition, participants had to learn the new mapping. However, the elongated RTs could not be entirely attributed to learning of the new stimulus-go association. Instead, the previous no-go association interfered greatly with learning the new go association. First, the RTs in the fourth block of the test phase were significantly longer than the RTs in the fourth block of the training phase (610 vs. 561 msec

Figure 3. Normalized MEP amplitude at (A) 100 msec and (B) 250 msec during the test phase in the consistent and inconsistent mapping groups. Automatic inhibition was evident as early as 100 msec as demonstrated by the lower MEP for go trials in the inconsistent mapping than in the consistent mapping condition (A). (C) and (D) show correlations between the amount of learning during the training phase and the MEP difference (go vs. no-go) during the test phase for the consistent and inconsistent mapping groups. Note that for the inconsistent group, go trials in the testing were no-go trial in the training, and vice versa.



for RTs in the fourth block of the test and training phase, respectively, $p < .05$). Second, the amount of learning within four blocks (RT difference between the fourth block and the first block) in the test phase was much smaller than that in the training phase (42 vs. 86 msec for test and training phase, respectively, $p = .05$). These results are consistent with the original findings reported in Verbruggen and Logan (2008a) and suggest that the old stimulus–stop associations interfered with responding in the test phase. Unlike the RT data, no significant main effect or interaction was found for $p(\text{respond} | \text{no-go})$ (see Figure 2B; all $ps > .6$).

MEPs

For each participant, the MEPs in each condition were normalized by dividing by each participant’s baseline MEPs (see Methods). The automatic inhibition account predicts that, at the early SOAs, there will be an interaction between trial type and mapping and, at the late SOAs, a main effect of Mapping and an interaction between Trial Type and Mapping. A full ANOVA of 2 Trial Types (go, no-go) \times 2 Mappings (consistent, inconsistent) \times 4 TMS Times (100, 150 200 250 msec) revealed a nearly significant three-way interaction, $F(3, 57) = 2.5$; $p = .07$. There was also a significant interaction between Trial Type and Mapping, and an interaction between Trial Type and TMS Times, $F(1, 19) = 5.65$, $p = .03$; $F(3, 57) = 3.62$, $p = .02$, respectively. No other effects were significant, $ps > .1$. These significant interactions indeed motivated further analyses, that is, separate ANOVAs of

2 Trial Types (go, no-go) \times 2 Mappings (consistent, inconsistent) at each TMS time.

At 100 msec (see Figure 3A), there was a significant interaction of Trial Type and Mapping $F(1, 19) = 5.4$, $p < .01$, and a significant main effect of Mapping, $F(1, 19) = 5.1$, $p < .04$. No other effect was significant ($ps > .4$). The interaction was due to lower MEPs for go trials than no-go trials in the inconsistent condition (one-tailed t test, $p = .05$, MEP ratio = .77 vs. .86 for go vs. no-go trials, respectively). However, we did not observe a reliable MEP difference between no-go and go trials in the consistent condition ($p > .5$).

At 250 msec (see Figure 3B), there was a significant interaction between trial type and mapping, $F(1, 19) = 6.6$, $p < .02$, and a significant main effect of trial type, $F(1, 19) = 5.4$, $p < .03$. The interaction was because of greater MEPs for go trials than for no-go trials in the consistent condition (one-tailed t test, $p < .01$); in the inconsistent condition, there was no reliable MEP difference ($p > .5$). Furthermore, the MEPs for go trials were larger in the consistent condition than in the inconsistent condition (MEP ratio = 1.1 vs. 0.8 for consistent vs. inconsistent, respectively, $p < .02$). No other effect was significant ($ps > .6$). None of the effects was significant at the SOAs of 150 and 200 msec ($ps > .5$). See Tables 2 and 3 for complete MEP data.

To further test whether the MEP suppression in the test phase was due to inhibition triggered by the retrieval of previously acquired stimulus–stop associations, we compared MEPs for go and no-go trials and examined whether this difference depended on the amount of learning during the training phase. For each participant, we

Table 2. Mean (*SE*) MEP Ratio for Each Condition in Experiment 1

Mapping	Trial Type	TMS Time			
		100	150	200	250
Consistent	go	0.98 (0.06)	0.90 (0.07)	0.96 (0.07)	1.08 (0.09)
	no-go	0.95 (0.05)	0.96 (0.08)	0.93 (0.06)	0.90 (0.08)
Inconsistent	go	0.77 (0.03)	0.84 (0.04)	0.85 (0.04)	0.84 (0.06)
	no-go	0.86 (0.05)	0.85 (0.06)	0.88 (0.07)	0.84 (0.05)

calculated the RT difference between the first four and the last four blocks of training and used this difference as an estimate of learning during the training phase. Note that we used RTs instead of $p(\text{respond}|\text{no-go})$ because the latter was already close to zero at the beginning of the training phase. We then calculated the MEP difference between go and no-go items in the test phase, and we analyzed the correlations between these two variables (behavioral learning and MEP difference) for the consistent and inconsistent mapping groups separately. There was a significant negative correlation in the inconsistent group (Pearson's $r = -.64; p < .05$). Figure 3D shows that the MEPs were *lower* for go trials than for no-go trials for participants who showed a large learning effect. This demonstrates that the MEP difference between go and no-go trials related to the amount of learning during training, supporting the automatic inhibition account. For the consistent group, there was no correlation between the amount of learning and the MEP difference between go and no-go trials (Pearson's $r = -.2, ns$; Figure 3C).

Finally, to verify that the FDI muscle was equally at rest before the TMS pulse in each condition, an ANOVA with 2 Mappings (consistent, inconsistent) \times 2 Trial Types (go, no-go) \times 5 TMS Times (-200, 100, 150, 200, 250 msec) was performed. No significant main effects or interactions were found (all $ps > .3$).

Discussion

Recent research has demonstrated that responding is slowed down for items that were previously associated

with stopping (Verbruggen & Logan, 2008a). Here we tested whether this behavioral observation could be explained by an underlying motor suppression mechanism. In the test phase of the paradigm, we delivered TMS at various time points after the presentation of the stimulus and found that MEPs were reduced for stimuli that were previously associated with stopping. This effect was observed as soon as 100 msec after the presentation of the stimulus (Figure 3A). Furthermore, we found that the motor suppression in the inconsistent condition was greater for those participants who learned more during the training phase (Figure 3D). These results strongly suggest that response slowing to items that were previously associated with stopping relates to a reinstatement of motor suppression.

Puzzlingly, however, we did not observe a reliable effect of "automatic inhibition" in the consistent condition (only a small numerical trend, see Figure 3A). First, MEPs for no-go trials at 100 msec were not lower than baseline nor were they lower than go trials (MEPs for go: 0.98 vs. no-go: 0.95, $ps > .5$). The automatic inhibition account predicted such difference because the no-go items were strongly associated with stopping. Furthermore, there was no correlation ($r = -.2; p > .5$) between behavioral learning and the MEP difference in the consistent condition (Figure 3C). We conjectured that the need for suppression of motor output disappears completely after substantial stimulus-stop practice and that in such situations "automatic inhibition" may no longer occur. Consistent with the idea that there is no longer suppression of motor output in the test phase of the consistent group, we found

Table 3. Mean (*SE*) MEP Ratio for Each Condition in Experiment 2

Learning Phase	Trial Type	TMS Time		
		100	200	300
Early (Blocks 1–3)	go	0.93 (0.05)	0.86 (0.05)	0.82 (0.05)
	no-go	0.89 (0.05)	0.79 (0.05)	0.77 (0.05)
Mid (Blocks 4–7)	go	0.93 (0.04)	0.83 (0.05)	0.80 (0.05)
	no-go	0.79 (0.04)	0.77 (0.06)	0.66 (0.05)
Late (Blocks 8–10)	go	0.87 (0.03)	0.77 (0.04)	0.79 (0.06)
	no-go	0.79 (0.03)	0.72 (0.03)	0.67 (0.05)

that, at an SOA of 250 msec, no-go MEP was still close to baseline whereas previous studies found a strong suppression (e.g., Yamanaka et al., 2002).

Thus, it was perhaps the case that automatic inhibition did develop to the no-go items in the consistent condition during learning but was no longer applied after substantial practice. One way to perform the task after a lot of practice is to simply ramp up the “going” response if the stimulus means go and not do this if it means no-go (see, e.g., Gomez, Ratcliff, & Perea, 2007, for a model of how this might occur); consequently, response suppression would no longer be exerted, and the “suppress” tag would be replaced by a “no response” tag. The corollary of this idea is that automatic inhibition or any stopping process, despite being learned, would not be initiated if no longer relevant (i.e., when there is no longer a “tendency to go” that has to be suppressed). This revised automatic inhibition is consistent with a recent finding by Verbruggen and Logan (2009) showing that primes such as the word “stop” do not have an effect on going when stopping is irrelevant to the task (e.g., in a choice reaction task in which subjects are informed that they can always go), but it does affect going when stopping is relevant. Thus, bottom-up or automatic inhibition primarily appears to influence performance in situations in which active suppression of incorrectly triggered responses is required.

This conjecture led us to a “revised automatic inhibition” hypothesis, which predicts that automatic inhibition emerges during training and then disappears with extensive practice, only to re-emerge in an “executive setting,” when, for example, the mapping changes and participants have to rely on active suppression again because old go items no longer require a response. To test this, we performed Experiment 2 in which TMS pulses were applied during 10 blocks of training. We expected that corticospinal excitability would be less for no-go than go in the midpoint of training (corresponding to the behavioral asymptote), but not at the beginning (not enough learning yet) nor at the end (suppression is no longer required). In other words, we predicted an interaction between learning phase (early, mid, late) and trial type (go/no-go).

EXPERIMENT 2

Methods

Participants

Fourteen volunteers (19–29 years old, mean age = 22 years, 7 men) from the University of California, San Diego, participated in this study for monetary compensation (\$15/hr). All participants were native English speakers, had normal or corrected-to-normal vision, passed the TMS safety screening, and reported no history of neurological impairment. Written informed consent was attained according to an institutional review board protocol. None of the volunteers had participated in Experiment 1.

Apparatus and Stimuli

The same apparatus was used as in Experiment 1. Only the smaller/larger categorization task was used in this experiment. Each word was presented 10 times in 10 blocks.

Procedure

All participants underwent 10 blocks of categorization task. The category–go/no-go mapping (e.g., smaller = go, and larger = no-go) was counterbalanced across participants. The trial progression was the same as the testing phase in Experiment 1, except for the TMS pulse time. During each trial, one TMS pulse was delivered at one of the time points (i.e., –200 [baseline], 100, 200, and 300 msec) after the onset of the stimulus word. In each block, TMS was delivered on 12 trials at –200 msec (i.e., before the word onset) and 20 trials at each of the other time points. The ordering of different TMS time points was randomized for each block and each participant.

TMS Procedure

The same TMS procedure was used as in Experiment 1.

Results

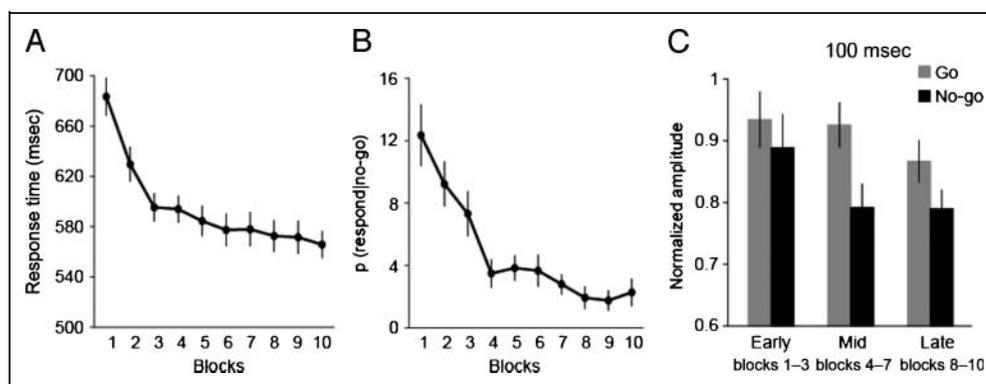
Behavioral Performance

Over the course of 10 blocks of training, RTs on go trials and probabilities of responding on no-go trials [$p(\text{respond} | \text{no-go})$] decreased (see Figure 4A and B), suggesting that participants learned the stimulus–go and stimulus–stop associations. There was a significant main effect of Training Block, $F(9, 117) = 28.28$ for RT; $F(9, 117) = 12.57$ for $p(\text{respond} | \text{no-go})$; both $ps < .0001$.

MEPs

To examine the MEPs at different learning phases, we binned the MEP data into three phases (*early*: Blocks 1–3, *middle*: Blocks 4–7, and *late*: Blocks 8–10) and conducted a 3 Learning Phases (early, middle, and late) \times 2 Trial Types (go, no-go) ANOVA separately for each TMS time (100, 200, 300 msec). On the basis of the findings of Experiment 1, we were especially interested in the earliest TMS time, that is, 100 msec. The ANOVA at 100 msec revealed a significant main effect of Trial Type, $F(1, 13) = 7.93$, $p < .02$, and a significant quadratic interaction between Learning Phase and Trial Type, $F(1, 13) = 5.93$, $p < .03$. The main effect of Phase was not significant, $F(2, 26) = 1.7$, $p > .1$. The quadratic interaction shows that the MEPs for no-go trials were significantly lower than go trials especially during the midphase of learning, but this effect was not evident yet during the early phase and was wearing off during the late learning phase (see Figure 4C).

Figure 4. Behavioral performance and MEP data in Experiment 2. (A) RTs for correct go trials. (B) Probability of responding during no-go trials or $p(\text{respond}|\text{no-go})$. Both behavioral measures indicated sufficient learning of stimulus go/no-go mapping in 10 blocks. (C) Normalized MEP amplitude at 100 msec poststimulus during the early (Blocks 1–3), mid (Blocks 4–7), and late (Blocks 8–10) learning phases for go and no-go trials. The quadratic interaction between go versus no-go and learning phase was significant ($p < .03$), suggesting that the suppression effect was most prominent during the midphase of learning.



The ANOVAs at 200 and 300 msec both revealed significant main effects of trial type, $F(1, 13) = 5.15, 16.17$, respectively, all $ps < .05$. No other main effects or interactions were significant, all $ps > .06$.

Finally, to provide a full view of the MEP data and ensure that we did not miss other interesting effects, we conducted a full ANOVA of 3 Learning Phases (early, middle, and late) \times 2 Trial Types (go, no-go) \times 3 TMS Times (100, 200, 300 msec). This ANOVA revealed a significant main effect of Learning Phase, $F(2, 26) = 3.7$, $p = .04$, trial type, $F(1, 13) = 15.48$, $p = .002$, and TMS Times, $F(2, 26) = 15.78$, $p < .001$. No other effects were significant, all $ps > .6$. However, the quadratic interaction between Trial Type and Learning Phase was marginally significant, $F(1, 13) = 1.9$, $p = .14$, which was related to the significant interaction at 100 msec reported above.

To verify that the FDI muscle was equally at rest before the TMS pulse in each condition, an ANOVA with 3 Learning Phases (early, middle, and late learning phase) \times 2 Trial Types (go, no-go) \times 3 TMS Times (100, 200, 300 msec) was performed. No significant main effects or interactions were found (all $ps > .5$).

In Experiment 2, we observed significantly reduced MEPs for no-go trials than go trials at merely 100 msec poststimulus, most prominently in the midphase of learning. Thus, in the early stage, the suppression was not well developed, likely due to not enough learning yet, and at the late stage, the effect was wearing off. This pattern of results is consistent with our earlier conjecture that automatic inhibition develops during learning when executive control is required but is not applied after substantial practice when the performance reaches asymptote and the need for active suppression of inappropriate responses disappears (i.e., the late learning stage of Experiment 2 and the consistent condition at the test phase of Experiment 1). The results here complement those of Experiment 1 and

demonstrate the development of automatic inhibition in the no-go trials during learning.

GENERAL DISCUSSION

The automatic inhibition hypothesis proposes that motor suppression is triggered by the retrieval of previously acquired stimulus–stop associations, leading to the slowing down of responses (Verbruggen & Logan, 2008a). Consistent with this, we observed, in Experiment 1, that MEPs at a mere 100 msec poststimulus were reduced for no-go_then_go trials compared with go_then_no-go trials in the inconsistent condition (Figure 3A). Thus, a history of stopping (no-go training with particular stimuli) can lead to subsequent quick motor suppression even if the participant has to now respond to those stimuli.

However, we did not see reliable motor suppression in the consistent mapping condition although automatic inhibition predicts lower MEPs for no-go_then_no-go than for go_then_go trials. To reconcile these findings, we supposed that automatic inhibition does develop in the consistent condition during learning but wears off as the performance asymptotes (or when less inhibitory control is needed as the suppression of incorrectly triggered motor responses is no longer required). Consistent with this new hypothesis, in Experiment 2, we demonstrated that MEPs for no-go trials during training were significantly reduced compared with go trials at a mere 100 msec poststimulus, but only during the midpoint of training (Blocks 4–7) around the time when performance started to asymptote. Taken together, the results from both experiments support a “revised automatic inhibition” hypothesis: Automatic inhibition does indeed involve motor suppression, but it only emerges in situations in which active suppression of motor output is required occasionally. Thus, “automatic inhibition” may be in the middle of a continuum between

effortful top-down response inhibition and no inhibition at all. When the need for overt response inhibition (an “executive setting”) disappears altogether, so does the manifestation of automatic inhibition. We suppose this is because when there is no executive setting, the “stopping network” is not active and so the “no-go tag” that was created during training cannot trigger the inhibitory system.

An alternative possibility that could explain the slowing down for inconsistent items in the test phase is a “pure conflict” account. On this view, there is conflict between the learned (but no longer relevant) plan to stop (no-go) and the new plan to go, and this leads to response slowing. When such conflict is detected, all motor output might be suppressed until the conflict is resolved (see, e.g., Frank, Samanta, Moustafa, & Sherman, 2007). This “mismatch” account predicts reduced MEPs (beneath baseline) for both go and no-go trials in the inconsistent condition, likely because of the caution for making correct responses. However, it does not predict a difference between go and no-go trials in the inconsistent condition because conflict should occur for both trial types. Although our data did show reduced MEPs (lower than baseline 1.0) for both go and no-go trials in the inconsistent condition, which is consistent with the first prediction of the conflict account, there was also a reliable difference in MEPs for no-go_then_go than go_then_no-go items at a mere 100 msec poststimulus. This latter result is not at all predicted by the pure conflict account. Another problem for the pure conflict account is that there was suppression at a mere 100 msec poststimulus. This seems too quick for a conflict process that occurs based on a mismatch of plans, which must require stimulus categorization and rule retrieval. For example, a recent ERP study (Randall & Smith, 2011) examined “conflict” in terms of the mismatch between cued action and target action (e.g., subjects were cued to go by left/right arrow and were presented with a no-go target) and found a difference in the N2 component. This suggested that a conflict process takes place around 200 msec after target onset (by our visual inspection, unlikely to be earlier than 150 msec). Assuming that after practice, the living/nonliving categorization became as easy as arrow left/right discrimination, we can assume that the conflict process would not have an effect on MEPs until 200 msec. Therefore, if the pure conflict account is true, any effect in the MEPs would not be detected until 200 msec or later.

By contrast, an effect within 100 msec does seem feasible for an automatic reinstatement of stopping because it does not need stimulus categorization. Consistent with this, a recent EEG study showed that, after a 30-min go/no-go training phase, there was modulation of evoked potentials to no-go stimuli (but not go stimuli) 61–104 msec after SOA (Manuel et al., 2010). The authors proposed that this early effect was consistent with a low-level form of inhibitory control. Thus, the currently observed motor suppression at a mere 100 msec after the stimulus is a plausible time frame for automatic inhibition to be in effect.

Yet, while a pure conflict account does not entirely explain our results, “conflict” between going and stopping does seem important for the expression of motor suppression, leading to a “revised automatic inhibition hypothesis. Our data argue that automatic inhibition emerges during training of go/no-go (in the midpoint, when “conflict” between go and no-go is putatively maximal), but it “disappears” later. We suppose that sometime toward the end of 10 blocks of training (Experiment 2) or 12 blocks of training (Experiment 1), the participant has had so much practice at the go/no-go discrimination that it becomes a matter of facilitating the response if it is go and not doing anything if it is no-go. Thus, a response tendency is no longer triggered on no-go trials after sufficient practice, so the need for both bottom-up and top-down inhibition in the task disappears altogether. When there is no longer any need for inhibition, retrieval of stimulus-stop associations may no longer need to suppress the motor output. As we pointed out above, this is consistent with a recent finding by Verbruggen and Logan (2009) showing that the word “stop” did not have an effect on going when stopping was completely “irrelevant” to the task. However, in the inconsistent mapping of Experiment 1, there was a response mapping reversal, consequently inhibition was occasionally required because responses previously associated with going were triggered incorrectly. Similarly, we assume that in the middle of training phase of Experiment 2, response inhibition was still required because of an incorrect tendency to respond on some trials. We propose that especially under such circumstances, “automatic inhibition” manifests itself.

The results of this study suggest that responses can be suppressed automatically by the retrieval of stimulus-stop associations. We suspect that the same cortical network mediating inhibitory control during deliberate stopping also implements automatic inhibition. Evidence for this was provided by Lenartowicz et al. (2011) whose fMRI study demonstrated that the rIFG was activated for stimuli that had previously been associated with stopping. Although the rIFG is part of a top-down inhibitory control network for deliberate stopping (e.g., Aron et al., 2007), activation of this region cannot be taken to unequivocally prove that inhibitory control is being generated during other conditions. However, our current findings with TMS provide strongly congruent evidence in favor of the idea that the rIFG activation in Lenartowicz et al. (2011) did reflect automatic inhibition. A related line of research by van Gaal and colleagues (van Gaal, Ridderinkhof, Scholte, & Lamme, 2010; van Gaal, Ridderinkhof, van den Wildenberg, & Lamme, 2009) has demonstrated that inhibition can be triggered by subliminally presented no-go primes. These primes slowed responding and activated the top-down inhibitory network (the rIFG and the presupplementary area). Furthermore, the stronger the activation in these regions, the greater the RT slowing because of the no-go primes (van Gaal et al., 2010). Thus, taken together, these results suggest that the inhibition network can be activated

in different ways: In a top-down fashion when an external, conscious signal is presented, or in a bottom-up fashion when stimulus-stop associations are retrieved, or when no-go signals are presented subliminally.

One note about our study is that we used the go/no-go paradigm instead of the modified version of the stop-signal paradigm in Lenartowicz et al. (2011). It is still a debatable question whether one can equate the inferred mechanisms involved in these two paradigms. Whereas some studies have shown that motor inhibition in the two paradigms engaged an overlapping neural network (e.g., Levy & Wagner, 2011; Chikazoe, et al., 2007), others have argued for dissociable circuits (e.g., Swick, Ashley, & Turken, 2011). Nonetheless, according to Verbruggen and Logan (2008a), automatic inhibition can be retrieved as long as there is consistent stimulus-stop mapping during training. Thus, our go/no-go paradigm provided a strong case for testing this hypothesis. Furthermore, in the modified stop-signal task, the number of critical trials for testing hypothesis is less than 50% of total trials, which makes the testing multiple TMS SOAs impractical.

How could bottom-up forms of inhibition contribute to goal-directed behavior in real life? Developing such forms of inhibition might be useful to modify motivated behaviors, as studies have shown that pairing “no-go” to a rewarding stimulus leads to a diminution of the motivated behavior (e.g., Houben, 2011; Wiers, Rinck, Kordts, Houben, & Strack, 2010; Veling & Aarts, 2009). In a behavioral study, Houben et al. (2011) instructed two groups of heavy drinking students to perform a task in which pictures of beer were paired with either go and no-go, respectively, and investigated how subsequent drinking behavior in these groups was modulated by the go/no-go training. They found that, a week later, participants in the beer/no-go condition reported a decrease in their weekly alcohol consumption compared with their pre-experiment consumption. These results suggest that learned stimulus-stop (e.g., beer/no-go) associations can effectively influence goal-directed behavior. Future therapeutic programs could leverage this approach for treating maladaptive behaviors in clinical disorders, which are characterized by poor self-control over urges and compulsions, such as obsessive-compulsive disorders, Tourette’s syndrome, or substance abuse. However, much work is needed to establish the neurocognitive mechanisms by which such therapeutic effects are mediated. Whereas in our study we provide evidence for automatic inhibition at the motor level in a semantic categorization task, it is uncertain whether this mechanism could apply in such circumstance as reduced drinking.

In conclusion, the current study shows that the retrieval of previously acquired stimulus-stop associations can automatically lead to motor suppression. Our MEP evidence of inhibition corroborates and extends previous findings in the literature, suggesting that response inhibition is controlled by both a top-down and a bottom-up mechanism. Finally, clinical treatments for impulse control

disorders could consider taking the approach of building up automatic inhibition to counteract specific motivated behaviors or overlearned motor tendencies.

Acknowledgments

This study was funded by the National Institutes of Health, National Institute on Drug Abuse (grant ROI-DA026452 to A. R. A.).

Reprint requests should be sent to Yu-Chin Chiu, 9500 Gilman Drive, #0109, University of California, San Diego, La Jolla, CA 92093, or via e-mail: chiu.yuchin@gmail.com.

REFERENCES

- Aron, A. R. (2011). From reactive to proactive and selective control: Developing a richer model for stopping inappropriate responses. *Biological Psychiatry*, *69*, e55–e68.
- Aron, A. R., Behrens, T. E., Smith, S., Frank, M. J., & Poldrack, R. A. (2007). Triangulating a cognitive control network using diffusion-weighted magnetic resonance imaging (MRI) and functional MRI. *Journal of Neuroscience*, *27*, 3743–3752.
- Arrington, C. M., & Logan, G. D. (2004). Episodic and semantic components of the compound-stimulus strategy in the explicit task-cuing procedure. *Memory & Cognition*, *32*, 965–978.
- Boehler, C. N., Appelbaum, L. G., Krebs, R. M., Hopf, J. M., & Woldorff, M. G. (2010). Pinning down response inhibition in the brain-conjunction analyses of the stop-signal task. *Neuroimage*, *52*, 1621–1632.
- Brainard, D. H. (1997). The psychophysics toolbox. *Spatial Vision*, *10*, 433–436.
- Chambers, C. D., Garavan, H., & Bellgrove, M. A. (2009). Insights into the neural basis of response inhibition from cognitive and clinical neuroscience. *Neuroscience and Biobehavioral Reviews*, *33*, 631–646.
- Chikazoe, J., Konishi, S., Asari, T., Jimura, K., & Miyashita, Y. (2007). Activation of right inferior frontal gyrus during response inhibition across response modalities. *Journal of Cognitive Neuroscience*, *19*, 69–80.
- Claffey, M. P., Sheldon, S., Stinear, C. M., Verbruggen, F., & Aron, A. R. (2010). Having a goal to stop action is associated with advance control of specific motor representations. *Neuropsychologia*, *48*, 541–548.
- Dodds, C. M., Morein-Zamir, S., & Robbins, T. W. (2011). Dissociating inhibition, attention, and response control in the frontoparietal network using functional magnetic resonance imaging. *Cerebral Cortex*, *21*, 1155–1165.
- Frank, M. J., Samanta, J., Moustafa, A. A., & Sherman, S. J. (2007). Hold your horses: Impulsivity, deep brain stimulation, and medication in parkinsonism. *Science*, *318*, 1309–1312.
- Gomez, P., Ratcliff, R., & Perea, M. (2007). A model of the go/no-go task. *Journal of Experimental Psychology: General*, *136*, 389–413.
- Hampshire, A., Chamberlain, S. R., Monti, M. M., Duncan, J., & Owen, A. M. (2010). The role of the right inferior frontal gyrus: Inhibition and attentional control. *Neuroimage*, *50*, 1313–1319.
- Houben, K. (2011). Overcoming the urge to splurge: Influencing eating behavior by manipulating inhibitory control. *Journal of Behavior Therapy and Experimental Psychiatry*, *42*, 384–388.
- Houben, K., Nederkoorn, C., Wiers, R. W., & Jansen, A. (2011). Resisting temptation: Decreasing alcohol-related affect and drinking behavior by training response inhibition. *Drug and Alcohol Dependence*, *116*, 132–136.

- Jahfari, S., Stinear, C. M., Claffey, M., Verbruggen, F., & Aron, A. R. (2010). Responding with restraint: What are the neurocognitive mechanisms? *Journal of Cognitive Neuroscience*, *22*, 1479–1492.
- Kučera, H., & Francis, W. N. (1967). *Computational analysis of present day American English*. Providence, RI: Brown University Press.
- Lenartowicz, A., Verbruggen, F., Logan, G. D., & Poldrack, R. A. (2011). Inhibition-related activation in the right inferior frontal gyrus in the absence of inhibitory cues. *Journal of Cognitive Neuroscience*, *23*, 3388–3399.
- Levy, B. J., & Wagner, A. D. (2011). Cognitive control and right ventrolateral prefrontal cortex: Reflexive reorienting, motor inhibition, and action updating. *Annals of the New York Academy of Sciences*, *1224*, 40–62.
- Manuel, A. L., Grivel, J., Bernasconi, F., Murray, M. M., & Spierer, L. (2010). Brain dynamics underlying training-induced improvement in suppressing inappropriate action. *The Journal of Neuroscience: The Official Journal of the Society for Neuroscience*, *30*, 13670–13678.
- Randall, W. M., & Smith, J. L. (2011). Conflict and inhibition in the cued-Go/NoGo task. *Clinical Neurophysiology*, *122*, 2400–2407.
- Rossini, P. M., Barker, A. T., Berardelli, A., Caramia, M. D., Caruso, G., Cracco, R. Q., et al. (1994). Non-invasive electrical and magnetic stimulation of the brain, spinal cord and roots: Basic principles and procedures for routine clinical application. Report of an IFCN committee. *Electroencephalography and Clinical Neurophysiology*, *91*, 79–92.
- Swick, D., Ashley, V., & Turken, U. (2011). Are the neural correlates of stopping and not going identical? Quantitative meta-analysis of two response inhibition tasks. *Neuroimage*, *56*, 1655–1665.
- van Gaal, S., Ridderinkhof, K. R., Scholte, H. S., & Lamme, V. A. F. (2010). Unconscious activation of the prefrontal no-go network. *The Journal of Neuroscience: The Official Journal of the Society for Neuroscience*, *30*, 4143–4150.
- van Gaal, S., Ridderinkhof, K. R., van den Wildenberg, W. P. M., & Lamme, V. A. F. (2009). Dissociating consciousness from inhibitory control: Evidence for unconsciously triggered response inhibition in the stop-signal task. *Journal of Experimental Psychology: Human Perception and Performance*, *35*, 1129–1139.
- Veling, H., & Aarts, H. (2009). Putting behavior on hold decreases reward value of need-instrumental objects outside of awareness. *Journal of Experimental Social Psychology*, *45*, 1020–1023.
- Verbruggen, F., & Logan, G. D. (2008a). Automatic and controlled response inhibition: Associative learning in the go/no-go and stop-signal paradigms. *Journal of Experimental Psychology: General*, *137*, 649–672.
- Verbruggen, F., & Logan, G. D. (2008b). Response inhibition in the stop-signal paradigm. *Trends in Cognitive Sciences*, *12*, 418–424.
- Verbruggen, F., & Logan, G. D. (2009). Automaticity of cognitive control: Goal priming in response-inhibition paradigms. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *35*, 1381–1388.
- Wiers, R. W., Rinck, M., Kordts, R., Houben, K., & Strack, F. (2010). Retraining automatic action-tendencies to approach alcohol in hazardous drinkers. *Addiction*, *105*, 279–287.
- Yamanaka, K., Kimura, T., Miyazaki, M., Kawashima, N., Nozaki, D., Nakazawa, K., et al. (2002). Human cortical activities during go/no go tasks with opposite motor control paradigms. *Experimental Brain Research*, *142*, 301–307.