The Practice of Going Helps Children to Stop: The Importance of Context Monitoring in Inhibitory Control

Nicolas Chevalier  
University of Edinburgh

Christopher H. Chatham  
Brown University

Yuko Munakata  
University of Colorado Boulder

How do we stop ourselves during ongoing action? Recent work implies that stopping per se is easy given sufficient monitoring of contextual cues signaling the need to change action. We test key implications of this idea for improving inhibitory control. Seven- to 9-year-old children practiced stopping an ongoing action or monitoring for cues that signaled the need to go again. Both groups subsequently showed better response inhibition in a Stop-Signal task than active controls, and practice monitoring yielded a dose-response relationship. When monitoring practice was optimized to occur while children engaged in responding, the greatest benefits were observed—even greater than from practicing stopping itself. These findings demonstrate the importance of monitoring processes in developing response inhibition and suggest promising new directions for interventions.

Keywords: cognitive control, inhibitory control, context monitoring, cognitive development, intervention

Supplemental materials: http://dx.doi.org/10.1037/a0035868.supp

While dining out, your phone buzzes. As you reach for it, your friend’s look reminds you this is rude. How do you manage to stop mid-reach? People must often stop themselves during ongoing actions—when making insensitive comments, reaching for tempting desserts, or sending e-mails without promised attachments. Poor response inhibition during childhood predicts reduced academic achievement, health, and income later in life (Blair & Razza, 2007; Moffitt et al., 2011). Deficits in response inhibition are associated with developmental disorders such as attention-deficit/hyperactivity disorder (ADHD) and autism (Alderson, Rapport, & Kofler, 2007; Luna, Doll, Hegedus, Minshew, & Sweeney, 2007). Thus, understanding and improving response inhibition and its development are of prime importance.

Most theories of response inhibition focus on cognitive and neural processes specialized for stopping per se. For example, interrupting the reach for a phone may depend most on processes specialized for inhibiting motor output. Such motoric stopping has been posited as a function of the right inferior frontal gyrus (rIFG; Aron, 2011; Aron & Poldrack, 2006; Chambers, Garavan, & Bellgrove, 2009). The rIFG is activated by the Stop-Signal task, in which participants perform a simple choice reaction task (e.g., pressing once on the same side as a target) unless a stop signal appears, in which case responses must be withheld (Verbruggen & Logan, 2008). The activation of rIFG is negatively correlated with the latency of response inhibition (Aron & Poldrack, 2006; Rubia, Smith, Brammer, & Taylor, 2003). Conversely, response inhibition is slowed when rIFG is damaged (Aron, Fletcher, Bullmore, Sahakian, & Robbins, 2003) or temporarily deactivated (Chambers et al., 2006).

However, recent work demonstrates the importance of monitoring the environment for contextual cues that signal the need to change action (Chatham et al., 2012; Dodds, Morein-Zamir, & Robbins, 2011; Hampshire, Chamberlain, Monti, Duncan, & Owen, 2010; Sharp et al., 2010). Interrupting the reach for a phone may depend most on successfully monitoring for relevant signals, like a friend’s disapproval. Evidence comes from tasks with different motoric demands but matched context-monitoring demands (e.g., Stop-Signal vs. a task that requires pressing a second time if a signal appears). Signatures of response inhibition (e.g., rIFG activation, relevant event-related potentials, and pupillometric measures) more closely track monitoring demands than motoric-
stopping demands, and behavioral measures of context-monitoring efficacy, but not stopping efficacy, predict both response inhibition performance and associated rIFG activation (Chatham et al., 2012). The present studies test a counterintuitive implication of this account for intervention: The practice of responding a second time if a signal appears should improve the subsequent ability to stop when the same signal appears because of the common demand to monitor for the signal and despite the mismatch in trained motor actions. By way of analogy, imagine trying to improve a child’s inhibition of hitting a playmate by first teaching that child to monitor for a relevant environmental signal, such as a disapproving adult, and to then hit the playmate twice. We tested this prediction in children, given greater trainability in childhood (Wass, Scerif, & Johnson, 2012), and suggestions that interventions in childhood are influential for outcomes (Diamond & Lee, 2011). We focused on 7- to 9-year-old children because they can engage control proactively (Chatham, Frank, & Munakata, 2009) and should therefore benefit from practice monitoring for signals before they appear.

Experiment 1

In Experiment 1, 7- to 9-year-old children practiced either motoric stopping, with an attempt to minimize context-monitoring demands, or context monitoring with no stopping demands. A third group practiced neither monitoring nor stopping but received similar experience otherwise. At test, all children performed a Stop-Signal task requiring monitoring for the same signal as the one used in Context-Monitoring practice, and withholding responses, as in Motoric-Stopping practice.

Method

Participants. Study participants included 88 7- to 9-year-old children in Experiment 1 (mean age = 8.4 years, SD = 0.5; 48 boys and 40 girls). Parental informed consent and children’s assent were obtained prior to participation. Children received small prizes, and parents received $5 compensation for travel.

Materials and procedure. All participants were tested individually in a 1-hr session. Children completed three experimental tasks while the parent sat behind them.

Initial assessment of response inhibition (10 min). Response inhibition was first assessed by digit identification accuracy on an Antisaccade task. Antisaccade and Stop-Signal performance correlate, and both tasks load onto the same latent factor in adulthood (Friedman et al., 2008; Miyake et al., 2000). (The Stop-Signal task was not used here to avoid interference between an initial experience with the task and later practice or test performance.) Children sat 60 cm from the screen and fixated a centrally presented cross. After a variable delay (ranging from 1,500 to 3,500 ms), a black square appeared on one side of the screen for 350, 375, or 400 ms, followed by a digit on the opposite side for 150 ms, which was then masked. Given this rapid timing, successful digit identification required inhibiting fixating the initial stimulus in order to quickly look in the opposite direction to identify the digit. After a warm-up series of 16 prosaccade trials and six antisaccade trials, children completed 3 × 18 test antisaccade trials.

Practice phase (30 min). Children were randomly assigned to the Motoric-Stopping condition (N = 29), Context-Monitoring condition (N = 27), or the Control condition (N = 32) (see Figure 1). Preliminary analyses showed that groups did not differ in age (MMotorizing = 8.4 years; MStopping = 8.3 years; MControl = 8.4 years), p = .878, sex distribution (Monitoring: 15 boys/12 girls; Stopping: 14 boys/15 girls; Control: 19 boys/13 girls), p = .680, or antisaccade score (MMotorizing = 0.44, SD = 0.21; MStopping = 0.39, SD = 0.20; MControl = 0.39, SD = 0.19), p = .538, suggesting they were comparable at study entry.

In the Motoric-Stopping condition, an airplane (target) appeared either on the left or right side of the screen on each trial. Children helped an air controller by pressing on the same side as the airplane to make it land. Feedback was provided in the form of an airplane being added to or removed from the bottom of the screen. After two demonstrations and 24 warm-up trials, Signal trials were introduced: After a variable delay, the background color turned dark gray, and clouds and lightning appeared below the target along with a thunder sound. The use of salient perceptual information as stop signals was intended to minimize context-monitoring demands by facilitating the detection of the need to stop. These trials required withholding responding because the storm made it too dangerous to land. The signal was presented after one of the following delays: 20%, 30%, 40% or 50% of each child’s mean reaction time (RT). These values were fixed throughout the practice phase to allow children to gauge improvement. To discourage slowing strategies, a response time limit of 1.5 × the child’s mean reaction time (computed based on the initial warm-up series of trials) was introduced on No-Signal trials.

The Context-Monitoring condition was identical to the Motoric-Stopping condition, except for the following changes. Children had to feed yellow bananas (targets) to an adult monkey. Feedback was provided by adding or removing a banana at the bottom of the screen. Critically, on Signal trials, the yellow banana turned brown (signal) after a variable delay. Children were instructed to quickly press again to make brown bananas go away because the monkey did not like those. The time limit for the first press was 1.1 × the child’s mean RT and 1.0 × the child’s mean RT for the second press. These values, set based on pilot participants, ensured that this condition was roughly as challenging and yielded the same amount of positive/negative feedback as the Motoric-Stopping condition.

A Control condition used the same cover story, targets, and signal as in the Context-Monitoring condition to ensure that any potential difference in test performance could not relate to differences in materials familiarity. However, children were instructed to press (once) on the same side as the banana on both No-Signal and Signal trials (i.e., irrespective of whether the banana remained yellow or turned brown). In addition, Picture trials, where a cartoon character was presented and children were instructed to just wait, kept children engaged and ensured that they did not build a stronger prepotency to respond. There were 27 No-Signal trials, nine Signal trials, and 12 Picture trials in each block.

Test phase (10 min). At test, all children completed the same Stop-Signal task (see Figure 1), which contained 24 warm-up trials without Signal trials, 24 warm-up trials with Signal trials, and three test blocks containing 36 No-Signal and 12 Signal trials each. Children had to feed yellow bananas to a baby monkey on No-Signal trials. On Signal trials, after a variable delay, the banana turned brown, in which case children were instructed to withhold

960 CHEVALIER, CHATHAM, AND MUNAKATA
responding to make it go away. Therefore, children had to monitor for banana color changes, as in the Context-Monitoring condition, but withhold responding, as in the Motoric-Stopping condition. We used the same signal across practice and test in the Context-Monitoring condition to maximize the transfer from monitoring for a signal; however, this design also maximizes the additional demand of overcoming a prepotent response to this signal at test. In contrast, signals differed between practice and test in the Motoric-Stopping condition to maximize benefits attributable to common stopping demands, rather than monitoring.

Inhibitory control at test was indexed by Stop-Signal Reaction Time (SSRT), an estimate of how efficiently responses can be stopped; faster SSRTs indicate better inhibitory control (Band, van der Molen, & Logan, 2003). Stop signal delay was adjusted following a staircase procedure (Verbruggen & Logan, 2009): The signal delay was increased by 50 ms after successfully inhibited responses and was decreased by 50 ms after errors on Signal Trials (unless the delay was already 0 ms).

**Bayesian model.** A parametric Bayesian method estimated the entire distribution of SSRTs using three parameters (see the online supplemental materials and Balota & Yap, 2011; Matzke, Dolan, Logan, Brown, & Wagenmakers, 2013; Matzke, Love, et al., 2013): $\mu$ (mu) for the mode (i.e., mean of the Gaussian component), $\sigma$ (sigma) for its dispersion (i.e., standard deviation of the Gaussian component), and $\tau$ (tau) for its positive skew (i.e., scale of its exponential component). These estimates of the entire dis-
tribution of SSRTs are more meaningful than the standard calculation of mean SSRT because they capture the fact that SSRTs vary across trials and they provide a more reliable measure with smaller numbers of trials (Matzke, Dolan, et al., 2013). We focus primarily on μ as the leading edge of the SSRT distribution and the modal time required for successful response inhibition; we consider implications of other measures in the General Discussion. Each estimate was analyzed separately for Signal and No-Signal trials with an analysis of variance (ANOVA) using condition as an independent variable and age as a covariate. Significant effects were further probed with Bonferroni-corrected contrasts.

Results and Discussion

Practice influenced inhibitory control, as indicated by all three estimates of SSRT distributions—μ: F(2, 84) = 13.25, p < .001, η² = .240; σ: F(2, 84) = 85.21, p < .001, η² = .670; and τ: F(2, 84) = 12.79, p < .001, η² = .234 (see Figure 2A). Children showed faster modal response inhibition after practicing context monitoring (μ = 247 ms) or motoric stopping (μ = 243 ms), relative to controls (μ = 288 ms), ps < .001. SSRTs were also less dispersed in the Context-Monitoring condition (σ = 46 ms) and, to a lesser extent, in the Motoric-Stopping condition (σ = 52 ms), relative to controls (σ = 65 ms), ps < .001, although the distribution of SSRT showed a longer tail in the Context-Monitoring condition (τ = 170 ms) and, to a lesser extent, in the Motoric-Stopping condition (τ = 137 ms) than in the Control condition (τ = 104 ms), ps < .037. In contrast, estimates for reaction times on No-Signal trials did not differ across conditions, ps < .368, suggesting the benefits of practice did not reflect greater motivation (see Figure 2B).

Figure 2. Performance in the Stop-Signal task (test) as a function of the practice condition. Context-Monitoring (N = 27) or Motoric-Stopping practice (N = 29) yielded better response inhibition (faster modal Stop-Signal Reaction Times [SSRTs], μ) in the Stop-Signal task, relative to controls (N = 32), with best performance after Simultaneous-Monitoring practice (N = 32). Across conditions, SSRT distributions varied in shape (A), whereas distributions of No-Signal reaction times (RTs) largely overlapped (B). Error bars indicate standard errors.
Table 1

Performance Improved During Practice in Context-Monitoring, Motoric-Stopping, and Simultaneous-Monitoring Conditions

<table>
<thead>
<tr>
<th>Condition</th>
<th>Part 1 Blocks 1–2</th>
<th>Part 2 Blocks 3–4</th>
<th>Part 3 Blocks 5–6</th>
<th>ANOVA</th>
</tr>
</thead>
<tbody>
<tr>
<td>Context-Monitoring (Experiment 1)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>No-Signal Accuracy</td>
<td>.90 (.07)</td>
<td>.92 (.05)</td>
<td>.91 (.06)</td>
<td>(F(2, 52) = 2.60, p = .084, \eta^2_p = .091)</td>
</tr>
<tr>
<td>No-Signal reaction time</td>
<td>527 ms (81)</td>
<td>510 ms (70)</td>
<td>499 ms (65)</td>
<td>(F(2, 52) = 6.28, p = .004, \eta^2_p = .194)</td>
</tr>
<tr>
<td>Signal Accuracy</td>
<td>.47 (.23)</td>
<td>.56 (.25)</td>
<td>.57 (.22)</td>
<td>(F(2, 52) = 10.66, p &lt; .001, \eta^2_p = .291)</td>
</tr>
<tr>
<td>Signal 1st press latency</td>
<td>604 ms (153)</td>
<td>581 ms (154)</td>
<td>557 ms (115)</td>
<td>(F(2, 52) = 5.10, p = .009, \eta^2_p = .164)</td>
</tr>
<tr>
<td>Signal 2nd press latency</td>
<td>929 ms (150)</td>
<td>925 ms (161)</td>
<td>914 ms (123)</td>
<td>(F(2, 52) = 0.58, p = .560, \eta^2_p = .022)</td>
</tr>
<tr>
<td>Motoric-Stopping (Experiment 1)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>No-Signal Accuracy</td>
<td>.82 (.15)</td>
<td>.83 (.15)</td>
<td>.84 (.15)</td>
<td>(F(2, 56) = 1.47, p = .237, \eta^2_p = .050)</td>
</tr>
<tr>
<td>No-Signal reaction time</td>
<td>544 ms (82)</td>
<td>535 ms (76)</td>
<td>524 ms (73)</td>
<td>(F(2, 56) = 8.66, p = .001, \eta^2_p = .236)</td>
</tr>
<tr>
<td>Signal Accuracy</td>
<td>.73 (.17)</td>
<td>.73 (.15)</td>
<td>.73 (.18)</td>
<td>(F(2, 56) = 0.01, p = .983, \eta^2_p = .001)</td>
</tr>
<tr>
<td>Control (Experiment 1)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>No-Signal Accuracy</td>
<td>.91 (.09)</td>
<td>.91 (.07)</td>
<td>.91 (.07)</td>
<td>(F(2, 62) = 0.18, p = .831, \eta^2_p = .006)</td>
</tr>
<tr>
<td>No-Signal reaction time</td>
<td>510 ms (71)</td>
<td>514 ms (71)</td>
<td>508 ms (71)</td>
<td>(F(2, 62) = 0.82, p = .443, \eta^2_p = .026)</td>
</tr>
<tr>
<td>Signal Accuracy</td>
<td>.91 (.07)</td>
<td>.91 (.07)</td>
<td>.91 (.09)</td>
<td>(F(2, 62) = 0.03, p = .966, \eta^2_p = .001)</td>
</tr>
<tr>
<td>Signal reaction time</td>
<td>510 ms (72)</td>
<td>514 ms (71)</td>
<td>520 ms (80)</td>
<td>(F(2, 62) = 0.89, p = .415, \eta^2_p = .028)</td>
</tr>
<tr>
<td>Simultaneous-Monitoring (Experiment 2)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>No-Signal Accuracy</td>
<td>.88 (.09)</td>
<td>.92 (.05)</td>
<td>.92 (.04)</td>
<td>(F(2, 62) = 9.50, p &lt; .001, \eta^2_p = .235)</td>
</tr>
<tr>
<td>No-Signal reaction time</td>
<td>519 ms (76)</td>
<td>501 ms (64)</td>
<td>485 ms (61)</td>
<td>(F(2, 62) = 17.04, p &lt; .001, \eta^2_p = .355)</td>
</tr>
<tr>
<td>Signal Accuracy</td>
<td>.47 (.24)</td>
<td>.56 (.23)</td>
<td>.66 (.19)</td>
<td>(F(2, 62) = 18.41, p &lt; .001, \eta^2_p = .373)</td>
</tr>
<tr>
<td>Signal 1st press latency</td>
<td>564 ms (188)</td>
<td>529 ms (94)</td>
<td>498 ms (79)</td>
<td>(F(2, 62) = 15.92, p &lt; .001, \eta^2_p = .339)</td>
</tr>
<tr>
<td>Signal 2nd press latency</td>
<td>906 ms (134)</td>
<td>871 ms (124)</td>
<td>845 ms (119)</td>
<td>(F(2, 62) = 12.85, p &lt; .001, \eta^2_p = .293)</td>
</tr>
</tbody>
</table>

Note. Standard deviations are presented in parentheses. Significant effects are bolded. ANOVA = analysis of variance.

As training-related improvement can predict cognitive control gains (e.g., Jaeggi, Buschkuehl, Jonides, & Shah, 2011), we examined whether performance improved over practice (see Table 1) and predicted response inhibition at test. Children responded increasingly fast in the Motoric-Stopping and Context-Monitoring conditions. In the Context-Monitoring condition only, the magnitude of this improvement (reduction in No-Signal RTs and first-press latency on Signal trials) predicted response inhibition (smaller \(\mu\), SSRT), \(r = -.586, p < .001\), and \(r = -.382, p = .049\), respectively. A linear regression, using the backward procedure, showed that only change in Signal RT significantly explained variance in \(\mu\), SSRT, \(\beta = -.85, t(25) = -3.61, p < .001\), adjusted \(R^2 = .317, F(1, 25) = 13.05, p < .001\) (see Figure 3).

This relation between improvement during practice and response inhibition at test suggests a way to increase the effectiveness of context-monitoring practice. Faster responding over the course of practice likely reflects children learning to monitor while engaged in responding, as opposed to monitoring for relevant signals and engaging in responding only after the signal has appeared or is unlikely to. Monitoring while engaged in action (i.e., after action initiation) is crucial for successfully inhibiting an ongoing action at test. Therefore, Context-Monitoring practice should be even more effective at improving response inhibition if children can be further encouraged during practice to monitor for the signal while engaged in responding.

Experiment 2

We built on the dose-response relationship between improvement during Context-Monitoring practice and response inhibition to optimize practice in Experiment 2. We tried to encourage children to learn to monitor while engaged in responding, by making the cue transient and increasing time pressure on the first press. We predicted that this new condition should reduce modal SSRTs at test relative to the conditions from Experiment 1.

Figure 3. Improvement during Context-Monitoring practice predicted modal Stop-Signal Reaction Time (SSRT). RT = reaction time.
Method

Participants, materials, and procedure. Thirty-two children participated in Experiment 2 (mean age = 8.3 years, SD = 0.5; 14 boys and 15 girls). They did not significantly differ in mean age, sex distribution (17 boys/15 girls), or antisaccade score ($M = 0.39$, $SD = 0.19$) from children in the other conditions, all $p > .685$. Materials and procedure were identical to Experiment 1. Participants completed the Simultaneous-Monitoring condition, which was similar to the Context-Monitoring condition, with two changes. First, time pressure for the first press on Signal trials was increased $(1.0 \times$ the child’s mean RT), whereas that on the second press was decreased $(1.2 \times$ the child’s mean RT) in order to encourage children to engage in responding quickly on the first press while keeping overall time pressure constant. Second, the banana flashed brown for 100 ms before turning back to yellow (instead of remaining brown) to encourage constant monitoring while engaged in responding.

Results and Discussion

ANOVAs including all four conditions showed effects of condition on all three SSRT estimates—$\mu$: $F(3, 115) = 35.18$, $p < .001$, $\eta^2_g = .479$; $\sigma$: $F(3, 115) = 88.10$, $p < .001$, $\eta^2_g = .697$; and $\tau$: $F(3, 115) = 31.42$, $p < .001$, $\eta^2_g = .450$ (see Figure 2A). Most importantly, children showed faster modal response inhibition after practicing simultaneous monitoring ($\mu = 201$ ms) relative to every other condition, $p < .001$ (see Figure 2A). The dispersion in the Simultaneous-Monitoring condition ($\sigma = 63$ ms) was greater than in the other practice conditions, $p < .001$, while not differing from controls ($p = .595$). The skew was also greater ($\tau = 211$ ms) than the other conditions, $p < .009$. Again, no differences in No-Signal reaction times reached significance, $p > .493$. Performance improved during Simultaneous-Monitoring practice but did not correlate with SSRT estimates, $p > .407$—consistent with the dose-response relationship arising from individual differences in context monitoring in Experiment 1 that we reduced with Experiment 2’s optimizations.

General Discussion

These studies demonstrate that children’s response inhibition can be improved through practice monitoring for contextual cues that signal the need to change action. The practice of context monitoring while engaged in action improved subsequent modal response inhibition more than practice with motoric stopping itself. Moreover, although we tried to minimize context-monitoring demands in the stopping condition, children may have nonetheless monitored for perceptually salient signals and thus benefited from practice with context monitoring. However, our targeted context-monitoring practice without motoric stopping was more effective. Furthermore, better modal response inhibition in the Simultaneous-Monitoring condition than the Context-Monitoring condition demonstrates that parametric manipulation of context-monitoring demands using otherwise equivalent stimuli shapes subsequent response inhibition.

Of note, after practice with stopping and, even more so, monitoring, advantages in the leading edge of the SSRT distribution ($\mu$ estimates) were accompanied by elevated tails ($\tau$ estimates). Thus, practice yields faster modal response inhibition, increasing the large proportion of trials that fall within the adult range of SSRTs (roughly 100–300 ms); however, it is also associated with an increase in the small proportion of extremely long SSRTs ($>500$ ms), which may reflect occasional neglect of the stopping goal (e.g., due to increased fatigue), especially if the cue has become associated with going again. Such tradeoffs could only be revealed by recent methodological advances in estimating SSRT (Matzke, Dolan, et al., 2013).

These findings reveal that monitoring for relevant contextual cues is critical to developing inhibitory control and suggest promising new directions for interventions. Previous attempts to train inhibition during childhood have shown limited success (reviewed in Diamond & Lee, 2011) but did not target context monitoring or assess effects on the entire distribution of SSRTs. Practice monitoring enhanced response inhibition, despite the mismatched motor demands between training and test. This mismatch isolated and tested the contributions of context monitoring but would not be necessary for real-world applications; children could practice monitoring for relevant contextual cues without learning an unhelpful response to them.

These broad theoretical and practical implications also motivate important next steps. Future work should determine whether context monitoring also contributes to response inhibition earlier in typical development and disinhibitory disorders, whether and how improvements in context monitoring drive developments in response inhibition, and how generalizable context-monitoring interventions are across situations and populations. Such work will inform an understanding of typical inhibitory control failures (like children hitting friends and adults checking phones), associated life outcomes, and clinical disorders linked to impaired inhibitory control.

References


