Improving Children's Response Inhibition: Effects of Active Computation, Passive Dissipation, or Additional Instructions and Reminders?

by

Jane E. Barker

B.A., Rice University, 2004

M.P.A., Indiana University Bloomington, 2008

A thesis submitted to the

Faculty of the Graduate School of the

University of Colorado in partial fulfillment

of the requirement for the degree of

Master of Arts

Department of Psychology and Neuroscience

2013

This thesis entitled:

Improving Children's Response Inhibition: Effects of Active Computation, Passive Dissipation, or Additional Instructions and Reminders?
by Jane E. Barker

has been approved for the Department of Psychology and Neuroscience

Yuko Munakata (Chair)

Akira Miyake

Tim Curran

Date: 12/4/2013

The final copy of this thesis has been examined by the signatories, and we find that both the content and the form meet acceptable presentation standards of scholarly work in the above mentioned discipline.

HRC protocol # 13-0090

Abstract

Jane Barker (M.A., Psychology and Neuroscience)

Improving Children's Response Inhibition: Effects of Active Computation, Passive Dissipation, or Additional Instructions and Reminders?

Directed by Dr. Yuko Munakata, Professor, Department of Psychology and Neuroscience, University of Colorado Boulder

Children notoriously struggle to inhibit incorrect, prepotent responses, but often improve if a delay is introduced before they can respond. Children may use delays to *actively compute* correct answers (Diamond et al., 2002). Alternatively, delays may improve child performance by allowing prepotent responses to *passively dissipate* (Simpson et al., 2012). A third, untested account posits that improvements previously attributed to delays may instead reflect the influence of additional instructions and reminders in delay conditions. The present study tests predictions arising from each account via a go/no-go box search task. Three-year-olds opened boxes to find stickers or left them shut, based on *go* and *no-go* cues. Each child completed one of four conditions crossing cue highlighting (hidden placement of the cue vs. visible placement of the cue and additional cue reminders) with delay period (responses allowed immediately versus responses allowed only after a delay). Additional instructions and reminders, rather than delays *per se*, drove improvements in child response inhibition, challenging both active computation and passive dissipation accounts.

Table of Contents

List of Tables

List of Figures

Improving Children's Response Inhibition: Effects of Active Computation, Passive Dissipation, or Additional Instructions and Reminders?

Three and 4-year-olds frequently fail to override prepotent response biases (e.g., yelling out an answer, or running after a ball into the street) with controlled, appropriate responses, even after repeated reminders from well-meaning adults ("Raise your hand before speaking!" "Look both ways before crossing the street!"). Young children's characteristically poor response inhibition has been investigated experimentally in a variety of tasks, including peg-tapping, where children are to tap once when the experimenter taps twice, and twice when the experimenter taps once (Diamond & Taylor, 1996; Luria, 1966); Simon tasks, where children respond to stimuli via spatially incongruent response options (e.g., left side key press for a right side image; Davidson et al., 2006; Van der Ven et al., 2012); and go/no-go, or stop-signal tasks, where children must occasionally withhold a typical response such as a button press (Dowsett & Livesey, 2000; Klenberg et al., 2001; Simpson & Riggs, 2006). In each task, young children repeatedly commit inhibitory errors, even after demonstrating adequate comprehension of task rules.

Curiously, although 3 and 4-year-olds struggle to withhold automatic responses under normal circumstances, they can overcome response prepotency when delays are introduced before they can respond. In the day-night Stroop task, children are asked to say "day" to a picture of a moon and stars, and "night" to a picture of a sun. Although 4-year-olds typically perform well during the first few trials of day-night Stroop, they often revert to mistakenly giving the opposite, prepotent response thereafter (Gerstadt et al., 1994; Diamond et al., 2002). Children make far fewer errors if an experimenter sings a ditty ("Think about the answer, don't tell me!") that prevents the child from responding immediately after seeing the image (Diamond et al., 2002). Delays have also yielded improvements in box search tasks (Figure 1), where children are instructed to either open boxes to find stickers or leave boxes shut based on cues placed on box lids (a no-go cue indicates a box is empty, or a go cue indicates a box contains stickers). When the experimenter reveals the cue and box simultaneously, children often open boxes they should have left shut. By contrast, children make considerably fewer errors when the experimenter reveals the box, then waits a few moments before placing the cue on the lid (Simpson & Riggs, 2007; Simpson et al., 2012).

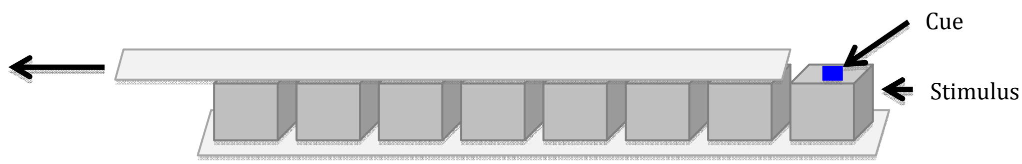

Figure 1. Schematic representation of box search task. The experimenter reveals each box sequentially by pulling a cardboard cover from right to left. In immediate conditions, the box (stimulus) and cue (blue square = *go*; red triangle = *no-go*) are revealed simultaneously. In delay conditions, the experimenter places the cue on the box lid ~2s after the box is uncovered.

How might delays benefit child response inhibition? One prominent explanation posits that children use delays to *actively compute* correct answers (Diamond et al., 2002). According to this account, children inhibit prepotent responses more effectively when, after viewing a stimulus, they pause for a moment to consider task rules. Because children often fail to pause when left to their own devices, they are less likely to make incorrect, prepotent responses when experimenters impose brief delays.

If children use delays to actively compute appropriate responses, one would expect that they must maintain focus on the task during the delay. This idea was tested using the box search task. Three-year-olds completed one of three conditions: immediate, delay, or delay-with-distraction (Simpson et al., 2012). In the delay-with-distraction condition, the experimenter revealed the box, then asked the child to guess which hand the cue was hidden in. After the child guessed (~ 2 s after the box was revealed), the experimenter placed the cue on the box lid. Critically, children demonstrated equivalently good performance in the delay and delay-with-distraction conditions, relative to the immediate condition, suggesting that delays confer benefits even when children cannot use the delay period to actively compute correct answers.

Evidence that delays benefit response inhibition even when children are distracted across the delay period has been interpreted as support for a *passive dissipation* account (Simpson et al., 2012; Simpson & Riggs, 2011). According to this account, prepotent and correct responses compete within a race model framework (Figure 2). Prepotent responses reach the response threshold before correct responses, then rapidly begin to weaken as correct responses begin their ascent. Passive dissipation theorists therefore argue that instituting a brief delay aids child response inhibition by giving the correct response time to strengthen (and eventually outcompete) the weakening prepotent response.

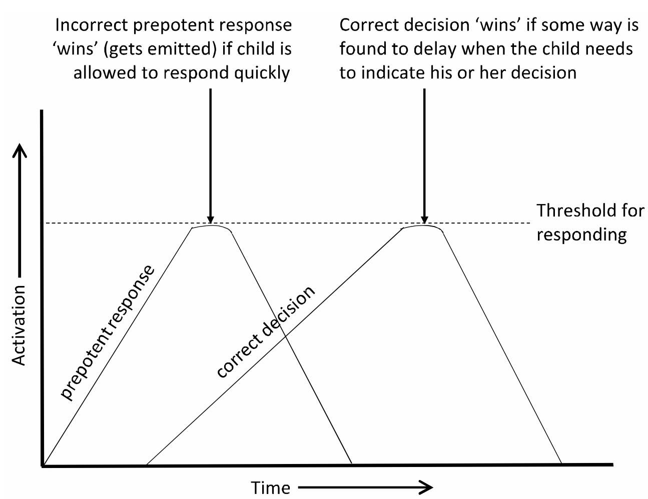

Figure 2. Passive Dissipation model (adapted from Simpson et al., 2012).

Although compelling, the passive dissipation account ignores a critical feature of prior inhibitory control tasks incorporating delays that may have benefitted child performance: in both day-night Stroop and box search tasks, delay conditions included additional instructions and reminders relative to immediate conditions. In the day-night Stroop task, the ditty instructed children to "Think about the answer" (Diamond et al., 2002), which may have cued participants to retrieve task-relevant information. In the box search task, the experimenter stated additional verbal instructions in the delay and delay-with-distraction conditions (e.g., "You mustn't open the box until I put the shape on top because only when the shape is on can you tell if there is a sticker inside"[1]; Table 1), the experimenter physically placed the cue on the box in view of the child in both the delay and delay-with-distraction conditions, drawing the child's attention to the cue, and potentially reminding them of the instructions, and the distractor task required that the child find the cue hidden in one of the experimenter's hands – a manipulation which again drew the child's attention to the cue (Simpson et al., 2012).

Additional verbal instructions and cue reminders in the delay conditions may have benefitted child performance in two distinct ways. First, additional verbal instructions in delay conditions may have increased the strength and completeness of behavioral goals and plans, or task representations, that children formed prior to test. Young children's representations can vary in strength (Morton & Munakata, 2002; Munakata, Morton & Yerys, 2003), so verbal reminders may have supplemented and reinforced initial understanding (for instance, by aiding child attempts to map the relationship between task cues (triangle, square) and actions (reach, don't reach)) (Towse, Lewis & Knowles, 2007; Zelazo et al., 2007; Muller et al., 2008). Thus, differences in performance between children in delay and immediate conditions may have been driven by differences in the strength of task representations available at test. Secondly, experimenter placement of the cue on the box surface in view of the child may have increased the salience of cues in delay conditions. Child inhibitory control reliably improves when cues are made more salient (e.g., Towse, Lewis & Knowles, 2007; Bluell & Montgomery, 2013; Martin-Rhee & Bialystok, 2008; Jeffrey, 1968).

Taken together, additional instructions and cue placement by the experimenter in delay

[1] Children also received additional reminders each time they opened a box during the delay period, prior to cue placement (detailed in Table 1).

conditions may have helped children to activate task-relevant goals and plans. For instance, cues and reminders may have decreased child *goal neglect* across the task, or "disregard of a task requirement, even though it has been understood" (Duncan et al., 1996; also Marcovitch et al., 2010; Marcovitch, Boseovski & Knapp, 2007). Similarly, experimenter placement of the cue in the moment before children began to reach may have helped children to activate task representations reactively, as needed (Chatham, Frank & Munakata, 2009; Munakata, Snyder, & Chatham, 2012; Blaye & Chevalier, 2011; Chevalier et al., in prep).

To test whether children's response inhibition is improved by delays *per se*, as would be predicted by the active computation and passive dissipation accounts, or instead by additional task instructions and increased cue salience, as would be predicted by accounts emphasizing the importance of each, the present study tested 3-year-old children in a box search task. Children were allowed to reach either immediately or after a delay, and cues were either "highlighted" (via additional rule repetitions and physical placement by the experimenter) or not.

Crossing delay and cue highlighting yielded four conditions: Immediate + Cue Highlighting, Delay + Cue Highlighting, Immediate without Cue Highlighting, and Delay without Cue Highlighting. Active computation and passive dissipation accounts yield a similar prediction: delays should improve children's response inhibition, either because they allow time for children to compute the appropriate task rule, or because they allow time for a prepotent response to passively dissipate. By contrast, accounts emphasizing the importance of task instructions and cue salience yield a competing prediction: cue highlighting should benefit response inhibition, regardless of whether or not the experimenter imposes a delay.

To supplement and extend analyses of child accuracy, relationships between task performance and condition were explored with respect to child go-trial reaction times (RTs) and spontaneous task-relevant speech. Previous box search studies have considered only accuracy (e.g., Livesey & Morgan, 1991; Simpson & Riggs, 2007; Simpson et al., 2012), so consideration of additional behavioral variables may capture important differences in performance across conditions.

Reaction time. Extant box search studies have not included assessments of child reaction times on go-trials (Simpson & Riggs, 2007; Simpson et al., 2012). However, at least one study has demonstrated a positive correlation between reaction time and inhibitory control in young children. In day-night Stroop, children who spontaneously delayed their responses after viewing the stimulus card demonstrated better performance than children who did not (Gerstadt et al., 1994). Thus, we predict that children who perform well on the box search task may spontaneously delay their responses in order to allow themselves time to recall and apply task rules.

Both the active computation and passive dissipation accounts suggest that spontaneous delays should benefit performance more in immediate, relative to delay conditions, since children in delay conditions have either (a) had time to prepare their responses during the delay, and therefore have less reason to delay after the cue has been placed (active computation account) or (b) passed through the phase of strongest prepotency, and therefore no longer have reason or need to delay (passive dissipation account; Figure 2).

Accounts emphasizing the importance of additional instructions and cue salience suggest that cue highlighting should enable faster responding, relative to conditions with no cue highlighting. For instance, if rule repetitions strengthen task representations, and stronger representations facilitate faster recall, children in conditions with additional instructions should be expected to demonstrate lower RTs than children in conditions with fewer rule repetitions. Likewise, if physical cue highlighting facilitates efficient, reactive retrieval of task representations, children may reach more quickly when their attention is drawn to the cue, relative to conditions where it is not.

Spontaneous task-relevant speech. Children often verbalize thoughts as a self-guiding regulatory aid when completing new or difficult tasks (Vygotsky, 1966; Meichenbaum & Goodman, 1969; Berk, 1992; Winsler et al., 2003; Kray, Eber, & Karbach, 2008). This 'private speech' may play a particularly important function in young children, who have difficulty maintaining task-relevant task rules in the face of distraction. However, not all young children employ private speech, and relatively little is known about when and how they may be prompted to adopt task-relevant speech spontaneously (Berk, 1992; Jones, Rothbart & Posner, 2003). Thus, exploratory analyses were undertaken to determine if children's spontaneous, task-relevant speech predicted fewer erroneous no-go responses in the box search task, as has been established in some tasks requiring inhibitory control (e.g., Kray, Eber, & Karbach, 2008), but not others (e.g., Jones, Rothbart & Posner, 2003). Additional analyses considered whether the frequency of task-relevant speech varied in Cue Highlight relative to No Cue Highlight conditions. Two hypotheses might explain such variance: cue highlighting may encourage children to spontaneously verbalize 'highlighted' elements of the task (e.g., shape or color). Alternatively, if cue highlighting facilitates more efficient retrieval of task representations, children may feel less need to engage self-regulatory speech, culminating in less frequent task-relevant speech in Cue Highlight conditions.

Method

Participants

One-hundred-fifty 3-year-olds (M_{age}= 3.52; range= [3.00, 3.99]; males=72) participated in the study, excluding 14 children dropped for non-participation (N=7), failure to comprehend study instructions (N=3), experimenter error (N=2), fussiness (N=1), and parent interference (N=1). Upon enrollment, each participant was assigned to a study condition, matching for age and gender ($N_{condition}$=30)[2]. All children spoke English as a first language. Participants were recruited from either

[2] Assignment to the first three conditions (Immediate + Cue Highlight, Immediate without Cue Highlight, and Delay with Cue Highlight) was random. Recruitment for the fourth (Delay without Cue Highlight) and fifth (Delay with Cue

a laboratory subject pool (N=100; 20 per condition) or from families attending the Children's Museum of Denver (N=50; 10 per condition). Museum participants were recruited via wall signs and badges worn by study personnel, and completed the experiment in a quiet museum area. Laboratory participants completed the experiment in a standard laboratory testing room. Informed consent was obtained for all participants, and parents were notified that they could cease participation at any point during the experiment. Parents of children recruited through the laboratory pool received $5 for travel expenses.

Materials

The box search apparatus replicated that described in Simpson and colleagues (2012) (Figure 1). For each of two sets of test boxes, eight white boxes (each 60 mm cubed; lids 65 mm wide x 65 mm long) were spaced equally along a white cardboard mounting strip (75 mm wide x 700 mm long). Across all trials, the 'go' cue was a blue square (40 mm each side), and the 'no-go' cue was a red triangle (40 mm each side). An additional set of four practice boxes (of identical dimensions to test boxes) were mounted to a shorter piece of cardboard (75 mm wide x 250 mm long). Separate strips of cardboard (75 mm wide by 800 mm for test boxes; 75 mm wide by 275 mm long in practice boxes) were used to cover test and practice boxes so that boxes remained hidden until revealed by the experimenter.

Procedure

Each child completed one of five conditions: Immediate without Cue Highlight, Immediate + Cue Highlight, Delay without Cue Highlight, Delay + Cue Highlight, or Delay + Cue Highlight, with screen. Each condition included demonstration, practice, and test phases. During the demonstration phase (which mirrored procedures described in Simpson et al., 2012), the

---

Highlight) conditions occurred later, and was partially random: approximately 75 percent of participants in the former and 50 percent of the participants in the latter condition were assigned randomly.

experimenter showed the child example 'go' and a 'no-go' boxes and explained that boxes with squares on top contained stickers, whereas boxes with a triangle on top did not. Children were told they could keep any stickers they found, and encouraged to retrieve the sticker from the 'go' demonstration box. Following the demonstration phase, the experimenter stated the condition instructions (Table 1) and presented the child with four practice boxes, revealing each box sequentially by moving the cardboard cover (Figure 1). Practice boxes alternated go and no-go trials. Standardized feedback was provided in each practice trial ("That's right! You should open that one, because there are stickers inside of it"; "That's right! You shouldn't open that one, because there are no stickers inside of it"; "Oops! Don't open that one, because there are no stickers inside of it"; "Oops, you should open that one, because there are stickers inside of it").

Table 1: Box search task condition instructions.

Task Phase	**Immediate w/o Cue Highlight (as in Simpson et al., 2012)**	**Delay + Cue Highlight (as in Simpson et al., 2012)**	**Delay w/o Cue Highlight**	**Immediate + Cue Highlight**	**Delay + Cue Highlight (with screen)**
Demonstration	Standard: "See these two boxes? One has a square and the other has a triangle on the top. If there is a square on the lid then there's a sticker inside the box, but if there is a triangle on top then there's no sticker in the box. So if you open the boxes with squares on top you'll win stickers! But leave the triangle boxes closed because there aren't any stickers in them."	Standard + "**Don't open the box until I put the shape on top, because only when the shape is on can you tell if there is a sticker inside."**	Standard + **"Now we're going to play a game where we look at this** [shows screen]**, then open boxes."**	Standard + "**Make sure you wait to open the box until you see the right shape!**"	Delay + Cue Highlight instructions + **"Now we're going to play a game where we look at this** [shows screen]**, then open boxes."**
Practice Feedback	Standard: **Correct go:** "That's right! You get a sticker when you open that one." **Correct no-go:** "That's right! You shouldn't open that one because there's no sticker inside." **Incorrect no-go:** "Oops! You shouldn't open that one because there's no sticker inside." **Incorrect go:** "Oops! You should open that one because there's a sticker inside."	Standard, **unless child reaches during the delay period, then: "Oops! Wait until I put the shape on top, because only when the shape is on top can you tell if there's a sticker inside!"**	Standard	Standard	Standard
Post-Practice	Standard: "Are you ready to find the stickers? Remember, open the boxes with squares on top because they have stickers inside, but leave the boxes with triangles on top closed because they are empty."	Standard + "**Don't open any boxes until I put a shape on top, because only when the shape is on top can you tell if there's a sticker inside."**	Standard	Standard + "**Make sure you wait to open the box until you see the right shape!**"	Standard + "**Don't open any boxes until I put a shape on top, because only when the shape is on top can you tell if there's a sticker inside."**
Test	None	**If child opens a box during the delay: "Don't open the box until you see the shape on top!"**	None	None	None

After children completed the practice trials, the experimenter removed the practice boxes, repeated the condition task instructions, and presented the child with the first of two sets of test boxes. Test boxes were arranged so that there were 8 boxes on each backing cardboard strip, for a total of 16 trials. Intermixed go and no-go cues (8 go, 8 no-go) were presented in the same pseudorandom order in all conditions. Children were given 3 s to initiate a reach towards the box[3]. If the child did not initiate a reach during the 3 s interval, the experimenter revealed the next box. The experimenter provided no feedback in test trials, except in cases when children erroneously opened boxes before the cue was placed in the Delay + Cue Highlight condition (detailed below). Experimenters recorded errors if, during the 3 s interval following cue presentation, a child failed to open a box on a 'go' trial, or erroneously opened a box on a no-go trial. Each session was also videotaped for later analysis. In laboratory sessions, cameras were placed behind the experimenter, and slightly to the right of the experimental table so that participant hands and the box apparatus were clearly visible. Occasionally, in museum settings, camera angles were modified so as to minimize videography of non-study participants.

Conditions

Immediate without Cue Highlight (as in Simpson et al., 2012). In the Immediate without Cue Highlight condition, the experimenter revealed the box and the cue simultaneously (Figure 3). Children were allowed to reach immediately, and received standard verbal task instructions (Table 1).

Delay + Cue Highlight (as in Simpson et al., 2012). In the Delay + Cue Highlight condition, the experimenter revealed the box, waited ~2.5 seconds, then placed the cue on top of the box lid

[3] If children became distracted during the task and stopped attending to the box apparatus, the experimenter said, "Are you ready?" and waited for the child to return gaze before revealing the next box in the sequence (occurred on 6% of all trials). Occasionally, children looked away from the box apparatus (e.g., down at stickers, towards a parent) as the experimenter began to reveal the box, before the cue was visible. In such instances the experimenter redirected the child's attention toward the box apparatus, then proceeded with the trial (occurred on <1% of all trials). These children were given 3 s from the time that they returned their attention to the box apparatus to initiate a reach.

(Figure 3). Children also received additional reminders during the demonstration and practice phases relative to the Immediate without Cue Highlight condition ("Remember, don't open the box until the shape is on top, because only if the shape is on top can you tell if there's a sticker inside") (Table 1). If a child opened a 'no-go' box during the delay in practice or test trials, the experimenter quickly reminded the child of the task instructions ("Wait until you see the shape!"), then placed the cue on the lid. Boxes opened during the delay period were counted as errors.

Immediate + Cue Highlight. In the Immediate + Cue Highlight condition, cues and boxes were revealed simultaneously, just as in the Immediate without Cue Highlight condition. However, two manipulations were introduced to mimic the cue highlighting incorporated in Simpson et al.'s (2012) delay condition. First, the experimenter quickly tapped the cue once with her right index finger as she revealed the box (Figure 3). This tap briefly highlighted the cue, but did not introduce a delay; children were allowed to reach immediately after the box and cue were revealed. Second, to replicate the additional verbal cue reminders given in Simpson et al.'s (2012) delay condition, children in the Immediate + Cue Highlight condition received additional verbal instructions related to the cue, relative to the Immediate without Cue Highlight condition. After each set of standard task instructions (Table 1), children were told, "Make sure you wait to open the box until you see the right shape!"

Delay without Cue Highlight. In the Delay without Cue Highlight condition, the experimenter revealed the box, then immediately placed a small colored screen (125 mm width by 150 mm length) between the box and the child so that the box lid was not visible from the child's perspective (Figure 3). After obscuring the child's view of the box with the screen, the experimenter placed the cue on the box lid, and then removed the screen so that the child could reach. The delay period between box reveal and screen removal lasted ~2.5 s, replicating the delay interval in the Delay + Cue

Highlight condition. Children received standard instructions during demonstration and practice, slightly modified to indicate that they would see a colored screen before opening the box (Table 1).

Delay + Cue Highlight, with screen. In the Delay + Cue Highlight + screen condition, modeled after the original Delay + Cue Highlight condition, the experimenter briefly (~1.5 s) placed a screen (identical to that used in the Delay without Cue Highlight condition) between the box and the child during the 2 s delay between box reveal and cue placement. Critically, the experimenter removed the screen before placing the cue so that the child could observe cue placement (Figure 3). Otherwise, the procedure was identical to that used in the standard Delay + Cue Highlight condition, barring a slight instructional modification (children were informed that they would see the screen during the task).

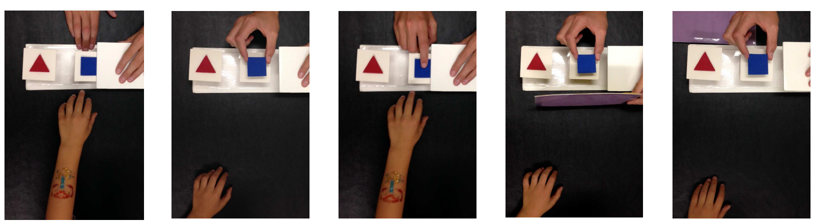

Figure 3: Visual comparison of box search conditions. From left to right: Immediate without Cue Highlight; Delay + Cue Highlight; Immediate + Cue Highlight; Delay without Cue Highlight; Delay + Cue Highlight, with screen (briefly introduced during delay, then placed behind boxes, as shown).

Coding

Three trained coders who were blind to all experimental hypotheses analyzed videos for 135 participants.[4] For each participant, each coder assessed experimenter script adherence, as well as child reaction times, verbal utterances, and attention.

[4] Videos for 15 participants were not coded because camera angle obscured reaching behavior (5 participants; all museum), equipment malfunctioned (6 participants), or parent refused taping (4 participants).

Go-trial RTs. Coders assessed go-trial reaction times by measuring, for each trial, the interval between the moment that cues were visible on top of the box (i.e., immediately in immediate conditions, and after placement in delay conditions) and the time that children began to lift the box lid. Go trial RTs were averaged across trials within participants to form composite go-trial RT scores.

Verbal Utterances. Coders analyzed all verbal utterances spoken by children during the task separately, in the absence of information identifying trial or condition. For each utterance, coders first established *relative time of utterance* (i.e., prior to cue visibility; prior to box opening, but after cue placement; or after box opening). Following this determination, responses were classified as either task-relevant; purely incidental (e.g., "I like this game"); or as incidental reactions to box contents (e.g., "I got a sticker!"). Task relevant utterances were further classified as containing one or more of the following: *reference to color* ("red" or "blue"); *reference to cue or shape* ("square", "triangle", "shape"); *reference to reward* ("There's a sticker in that one"); or *reference to action* (e.g., 'I'm going to open it'). References to each dimension were summed for each utterance to create a *trial-level task-relevant language score* (e.g., a single utterance referencing three dimensions – color, cue, and reward – would receive a score of 3). Scores for each utterance were then summed within participants to generate *subject-level task-relevant language scores.*

In a final, separate round of coding, utterances were judged as either restatements of task rules (e.g., "No, I shouldn't open that one because there is a triangle on top"; "No, the triangle means that the box is empty") or other forms of task-relevant verbalizations (e.g., "that one's blue"; "I'm going to open it"). Rule restatements falling into the former category were summed within participants to create *rule restatement scores.* Incidental language, reactions to box contents, and task-relevant language were summed within participants to create a *subject-level total language scores.*

Attention. Coders established child attention on each trial by assessing child gaze at the time that the experimenter revealed the trial box. Coders judged where children were attending (e.g., looking at trial box, looking at stickers, looking left or right, looking at another box), and subsequently categorized each judgment as either 'attending' (score=1; awarded if children were looking at trial box) or 'not attending' (score=0; awarded if children were not looking at trial box). Trial-by-trial attention scores were summed to create an aggregate *attention score* for each subject.

Results

Data were analyzed via a generalized linear mixed effects model (GLMM), a common, preferred alternative to ANOVA for analyzing repeated, within-subject binary outcome variables (e.g., 'correct' versus 'incorrect' trial performance) (Jaeger, 2008; Bolker et al., 2009; Laubrock et al., 2007). GLMMs estimate outcomes as a linear combination of fixed (e.g., condition, trial type) and random effects (e.g., within-subject variance). All analyses were performed using the R statistics package lme4 version 1.0-4 (Bates, 2007; Bates, 2013) using Laplace approximation for maximum likelihood estimation and a logit link function (standard procedures for binary outcome variables; see Bolker et al., 2009).

To determine how trial type and condition influenced the probability of correct performance on individual trials, performance (correct=1, incorrect=0) was predicted using trial type (go, no-go), condition (delay yes/no, cue highlight yes/no), and participant age as fixed effects. Subject intercepts were modeled as random effects.[5] Figure A-1 (Appendix A) illustrates model fit across each condition.

[5] Formally, the model was specified as Performance ~ TrialType + Delay + CueHighlight + Age_C + (1 | SubjectID), where Delay and Cue Highlight represent contrast-coded predictors indicating condition status (cue highlight versus no cue highlight; delay versus no delay), and age indicates child age in days (mean-centered). In the basic model, Trial Type was contrast coded (go = 1, no-go = -1) to generate an intercept representing average task performance across both trial

On average across conditions, children were more accurate in go trials (98%) than in no-go trials (75%) (b = 2.56; z = 14.74; p < .001; Table A-1). Younger children made more errors than older children across all trials (z=3.13; p < .002), but were not more likely to err in no-go trials than go trials (p > .9). Gender did not significantly predict performance (p > .3).

No-Go Trial Performance

As predicted by accounts emphasizing the importance of cue highlighting, children erroneously opened boxes less often when cues were highlighted, regardless of length of delay: children in Cue Highlight conditions ($M_{Imm+CueHL}$= 81%; $M_{Delay+CueHL}$= 83%; $M_{Delay+CueHLscreen}$= 80%) demonstrated better no-go accuracy than children in No Cue Highlight conditions ($M_{Imm-CueHL}$= 63%; $M_{Delay-CueHL}$= 68%) (Figure 4). A typical child in a Cue Highlight condition was 1.8 times as likely to perform a given no-go trial correctly as a child in a No Cue Highlight condition (b = .597; z = 9.61; p < .001; Table A-2).

Delays did not benefit response inhibition independent of cue highlighting. Children performed no better in Delay conditions, relative to Immediate conditions (p > .9; Table A-1). Planned group comparisons also demonstrated that delays alone did not benefit performance within levels of cue highlighting (Delay + Cue Highlight - Immediate + Cue Highlight contrast: p > .65; Delay without Cue Highlight - Immediate without Cue Highlight contrast: p >. 9).

As predicted, highlighting the cue *did* improve children's response inhibition. Children in the Immediate + Cue Highlight condition made fewer errors than children in No Cue Highlight conditions, both within and across levels of delay (Immediate without Cue Highlight contrast: b = 2.43; z = 1.97; p < .05; Delay without Cue Highlight contrast: b = 2.40; z = 2.00; p <. 05).

types. In subsequent iterations (Appendix A), Trial Type was dummy coded to generate an intercept representing average no-go trial (no-go = 0, go = 1) or go trial (no-go = 1, go = 0) performance, and condition-level interactions were added.

Corresponding contrasts between the Delay + Cue Highlight condition and No Cue Highlight conditions were marginally significant (Delay without Cue Highlight contrast: b = 1.81; z = 1.57; $p < .12$; Immediate without Cue Highlight contrast: b = 1.84; z = 1.56; $p < .12$). The observed cue highlight performance advantage was not improved (or attenuated) when cue highlighting was combined with a delay (the interaction of Cue Highlight x Delay conditions was insignificant; $p > .6$). Thus, planned contrasts also support the finding that additional instructions and reminders, rather than delays, benefit child performance on the box search task.

Go Trial Performance

As expected, performance in go trials was quite good across all conditions (by condition: $M_{Imm+CueHL}$= 99%; $M_{Delay+CueHL}$= 99%; $M_{Imm-CueHL}$= 100%; $M_{Delay-CueHL}$= 97%; $M_{Delay+CueHLscreen}$= 97%; Figure 5). Children in Delay conditions demonstrated slightly worse go trial performance than children in Immediate conditions (b = -1.71; z = -2.37; $p < .02$). However, this effect was quite small. There was no difference in go-trial performance between children in Cue Highlight and No Cue Highlight conditions ($p > .3$).

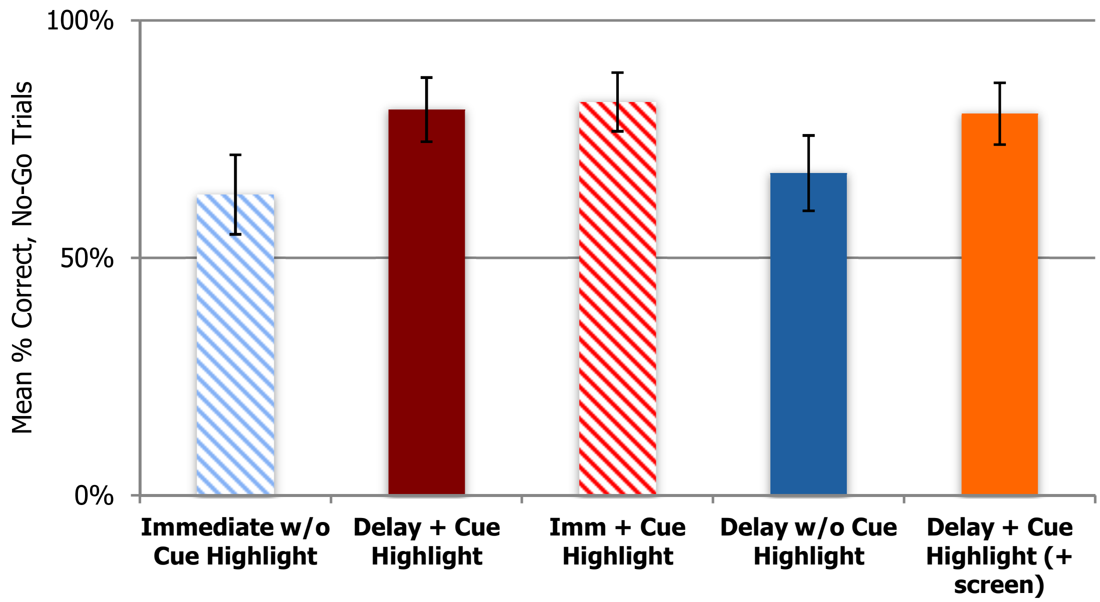

Figure 4: Average no-go trial accuracy by condition. Children in Cue Highlight conditions (red-orange shading) demonstrated significantly better performance that children in No Cue Highlight conditions (blue shading). Average performance in Delay conditions (solid) did not differ from average performance in Immediate conditions (striped). Error bars indicate 95% confidence intervals.

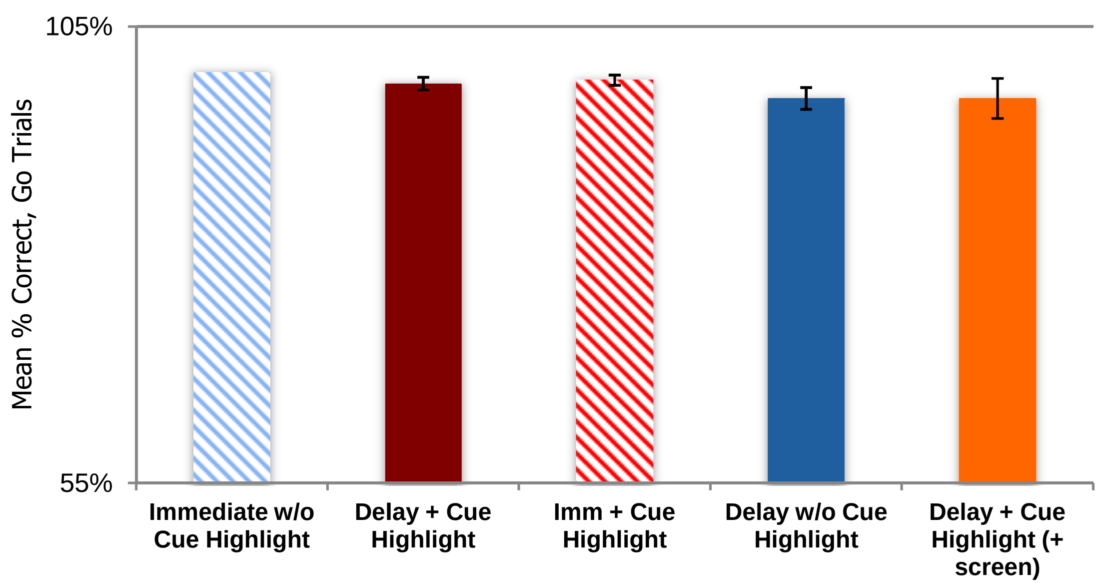

Figure 5: Average go trial accuracy by condition. Children in the Immediate without Cue Highlight, Immediate + Cue Highlight, and Delay + Cue Highlight demonstrated perfect or nearly perfect performance on go trials. By contrast, children in the Delay without Cue Highlight and Delay + Cue Highlight (with screen) conditions occasionally erred. Error bars indicate 95% confidence intervals.

Effect of Screen in Delay Conditions

To test whether any poor performance in the Delay without Cue Highlight condition was driven by the introduction of the screen, we introduced the same screen in a modified version of the Delay + Cue Highlight condition. As expected, children in the Delay + Cue Highlight + screen condition demonstrated marginally better performance than children in the Delay without Cue Highlight condition (included screen), outperforming children on no-go trials (80% versus 68% accuracy; $b = 1.82$; $z = 2.85$; $p < .07$), and demonstrating identical performance in go trials (97%). Children in the screen condition did not reliably differ in performance from children in the original Delay + Cue Highlight condition ($p > .7$), suggesting that the screen's introduction did not drive performance decrements.

Analysis of Attention By Condition

Separate analyses were undertaken to ensure that differences in performance across conditions were not driven by differences in participant attention. Summed participant attention scores (coded as 1 = attending or 0 = not attending for each trial; see coding section for details) were submitted to the basic GLMM used in preceding sections. Although child attention scores predicted trial by trial performance ($p < .02$; Table A-3), scores did not vary by group (Delay contrast $p > .2$; Cue Highlight contrast $p > .9$), suggesting that differences in child attention across conditions did not drive observed performance differences. Interactions by condition also failed to reach significance (p's $> .3$)

Effect of Experimental Setting

Additional tests were performed to assess whether distractions inherent to the museum environment influenced participant performance. Children who completed the experimental session in the museum demonstrated poorer performance, across conditions, than children who completed the session in the laboratory ($b = -1.71$; $z = -2.94$; $p < .02$; Table A-4). However, there were no significant interactions between experimental setting and Cue Highlight ($p > .3$) or Delay condition ($p > .8$) performance. Pairwise analyses assessing the effect of museum participation on performance within each condition also failed to reach significance (p's > .12), suggesting that the inclusion of museum participants did not drive observed condition differences.

Secondary Analyses

Task-Relevant Language

Within each trial, child use of task-relevant language marginally predicted better performance ($b = .215$; $z = 1.93$; $p < .055$) (Table A-5). This relationship was also observed across trials: summed task-relevant language scores positively correlated with overall no-go trial performance ($F = 4.84$; $p < .03$; Figure 6). By contrast, overall language scores, which captured both task-relevant and incidental language, were not predictive of performance ($p > .3$), suggesting task-relevant language, rather than verbalization in general, predicted inhibitory control. There were no significant interactions between child language and condition, either in the trial-by-trial model (p's > .8), or in the model summing total task-relevant language across all trials (p's > .4). Child rule restatements were not predictive over and above task-relevant language scores ($p > .8$).

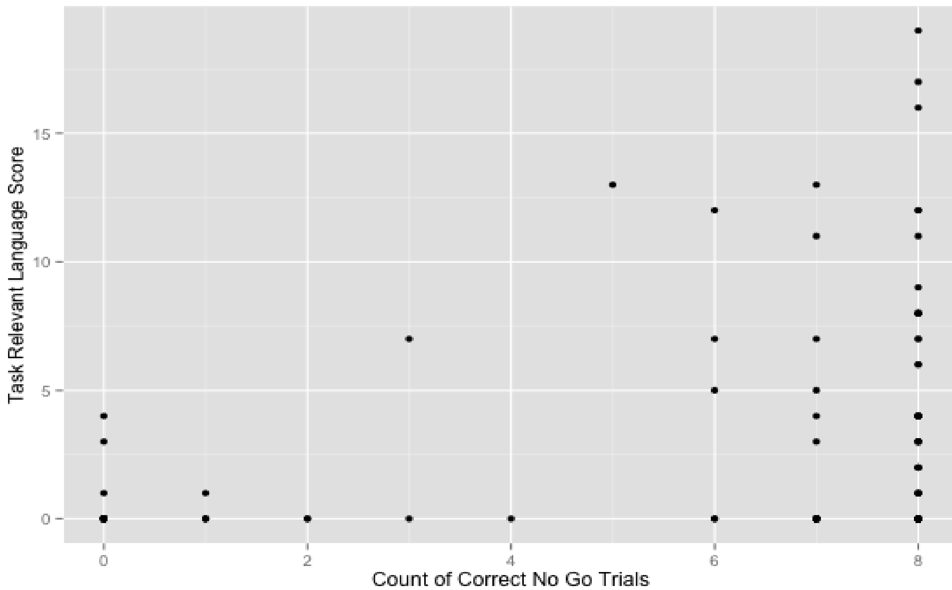

Figure 6: Relationship between task relevant language score (summed task-relevant language across all trials) and no-go trial performance. Children who used more task-relevant language were less likely to erroneously open no-go boxes (*F*= 4.84; p<.03).

Go Trial Reaction Times

For reaction time analyses, log transformed mean go-trial reaction times for each participant were submitted to a model including participant's total correct no-go trials, total attention score, delay condition (yes, no), and cue highlight condition (yes, no) as predictors[6].

In keeping with a previous finding (Gerstadt et al., 1994), children in the present study who reached more slowly on go trials made fewer errors in no-go trials ($b = .05$; $F(134) = 8.49$; $p < .005$). Visual inspection of RT plots suggested that outliers might have driven the observed relationship between no go trial performance and go trial RT (Figure 7; Appendix B, Figure B-1). Outliers were subsequently identified and replaced with the go trial RT grand mean (procedure described in

[6] Model specification: Log(MeanGoTrialRT) = $\beta_0 + \beta_1$NoGoCorrect + β_2Delay + β_3CueHL + β_4Age + β_5AttentionScore, where NoGoCorrect indicates subject's total correct no-go trials, Attention Score indicates subject's total attention score, and Delay and Cue Highlight (CueHL) represent contrast-coded categorical predictors (interactions omitted). Attention was included as a predictor to increase model power.

Appendix B, Figure B-1). After correcting for the influence of outliers, better performance on no-go trials continued to predict slower reaching in go trials a (b = .037; $F(134) = 5.76$; $p < .02$; Table A-6).

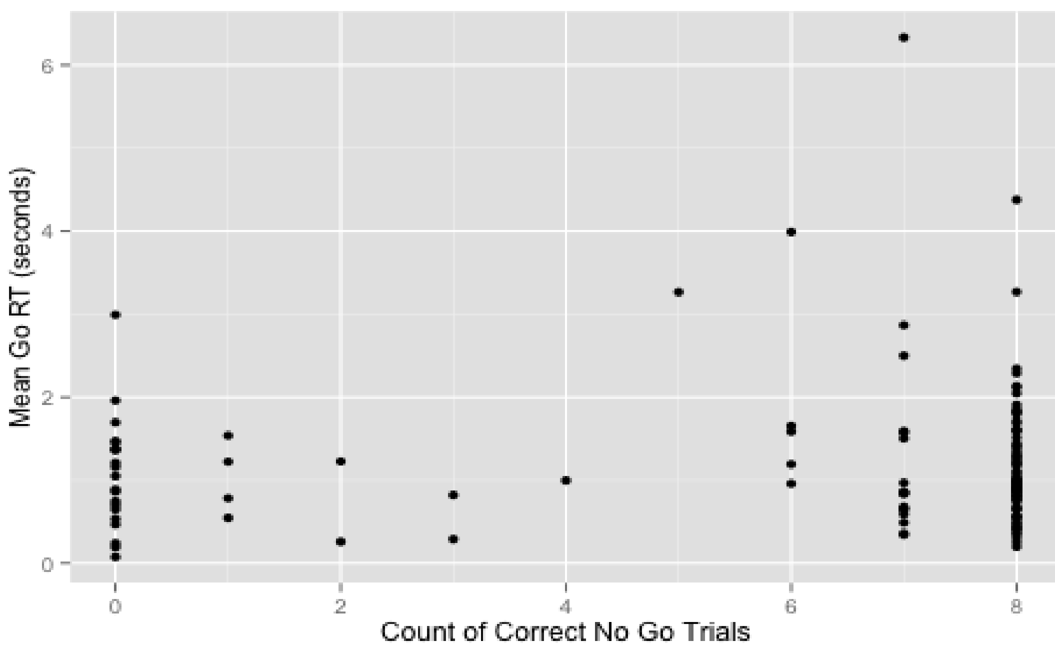

Figure 7: Relationship between mean go-trial RT and no-go trial performance. Children with better no-go trial performance were slower to reach on go trials, on average (b= .05; F(134)=8.49; p<.005).

Children in Delay conditions tended to reach more slowly, on average, than children in Immediate conditions (average Delay RT (corrected for outliers) = 1.41 s; b=.187; $F(134)=15.76$; $p < .001$; Table A-6; Figure 8), suggesting that they did not use delays to recall task-relevant information and prepare for future action, as would be predicted by active computation and passive dissipation accounts. However, the introduction of the screen in two out of the three delay conditions may have simply slowed children down. The average go-trial RT in Delay conditions including a screen was 1.57 s, significantly slower than the average RT in the Delay + Cue Highlight condition (.98 s), which did not include a screen (b = .412; $F(134) = 8.69$; $p < .004$).

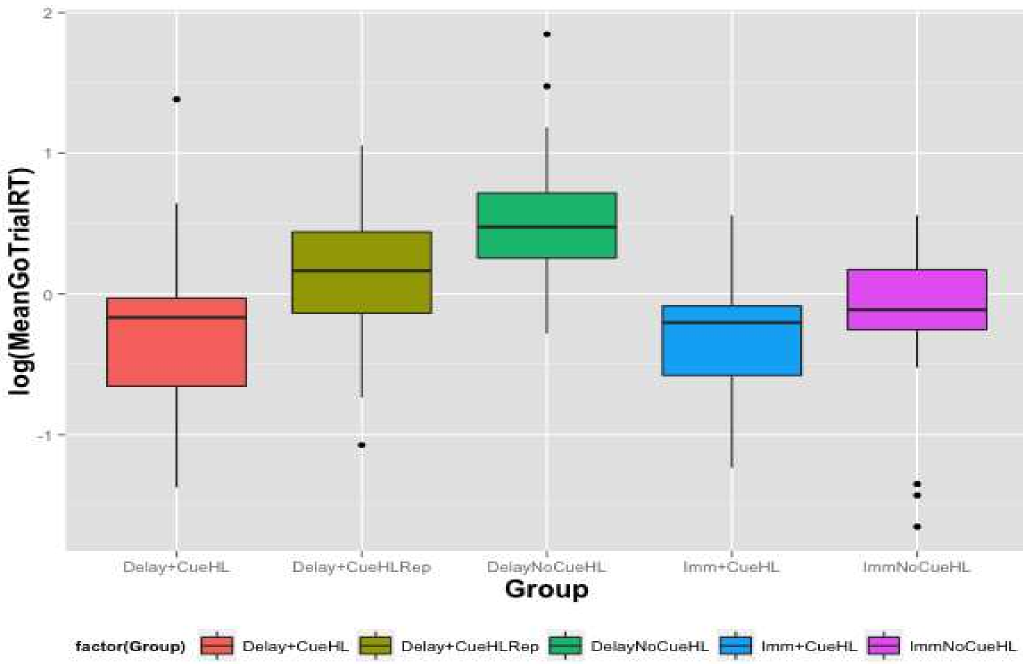

Figure 8: Logged mean go-trial RT performance by condition. Children in Delay conditions tended to reach more slowly, on average, than children in Immediate conditions (average Delay RT (corrected for outliers) = 1.41 s; b=.187; $F(134)$=15.76; p < .001), and children in Cue Highlight conditions reached more quickly than children in No Cue Highlight conditions (average Cue Highlight RT (corrected for outliers) = .99 s; b= -.21; $F(134)$=19.89; p < .001). Error bars represent ~95% confidence intervals.

Cue highlighting predicted a performance pattern opposite to that observed in Delay conditions: children in Cue Highlight conditions reached more quickly on go-trials than children in No Cue Highlight conditions, regardless of delay (average Cue Highlight RT (corrected for outliers) = .99 s; b= -.21; $F(134)$=19.89; p < .001). To account for differences in reaching resulting from the introduction of screens, we considered Cue Highlight-No Cue Highlight group contrasts in the presence and absence of screens. The effect of cue highlighting on go-trial reaction times was mixed in screen-less conditions: children were marginally faster to reach in the Immediate + Cue Highlight condition (average RT = .80 s) relative to the Immediate without Cue Highlight condition (p < .07), but demonstrated equivalent performance in the Delay + Cue Highlight and Immediate without Cue Highlight conditions (p > .2). In screen conditions, children were significantly faster to reach in the

Delay + Cue Highlight condition, relative to the Delay without Cue Highlight condition (b = .411, $F(134)=8.69$, $p < .004$). Thus, present findings suggest that cue highlighting may support faster responding in the box search task.

Discussion

The present study investigated why young children are better at inhibiting prepotent responses if a delay is introduced before they can respond, and whether performance benefits attributed to delays in prior work might instead be explained by additional instructions and reminders in delay conditions. Three-year-old children demonstrated reliably better inhibitory control in a box search task when the experimenter highlighted the cue, regardless of whether a delay was imposed. This result does not support a critical role of delays in response inhibition, and thus strongly challenges active computation and passive dissipation explanations of successful inhibitory control in young children.

Additional verbal instructions and physical cue highlighting may have contributed to child performance in separate but complementary ways. Instructional reminders in Cue Highlight conditions may have strengthened children's initial task representations, providing support for subsequent goal-directed responding. Later in the task, experimenter placement (or tapping) of the cue most likely drew children's gaze towards the cue, reminding them of its importance, and encouraging reactivation of goals and plans strengthened via earlier rule repetitions.

Although we have emphasized the joint influence of verbal and physical reminders, it is possible that physical cue highlighting could yield similar performance benefits even in the absence of additional task instructions. Two aspects of the task suggest that physical cue highlighting may have more strongly influenced inhibitory control than additional verbal reminders. First, experimenters reminded children to avoid reaching in the delay period (or, in the Immediate + Cue

Highlight condition, to attend to the cue), but did not include additional repetitions of each cue's meaning (e.g., "if you see a red triangle, you should leave the box closed"), which would have more strongly reinforced existing task representations. Secondly, physically highlighting the cue directly prior to the response would seem to be a much more salient reminder of the cue's meaning than instructions delivered several minutes before. In any case, future studies might easily investigate the relative influence of additional instructions in the box search task by varying pretest task instructions in otherwise equivalent box search conditions.

On average, children in Cue Highlight conditions opened boxes more quickly than children in No Cue Highlight conditions, suggesting that cue highlighting may support faster responding. There are several possible explanations for this finding. Physical highlighting may have increased the efficiency with which children retrieved task relevant information in the moment before they were required to respond. Secondly, additional instructions and reminders may have facilitated response preparation by aiding children's attempts to maintain task rules. A third, unrelated explanation is that experimenter 'prompting' (by touching the cue in the Immediate + Cue Highlight condition, or placing the cues in Delay + Cue Highlight conditions) may have encouraged children to speed their responses. Replication of the present study using a computerized version of the box search task could help to address the third explanation by reducing the experimenter's role in highlighting cues (e.g., by utilizing visual manipulations of the cue salience).

Adoption of a computerized version of the box search task might also help to rectify issues introduced by the inclusion of a screen in delay conditions. In the present study, children may have chosen to delay their reaches until the screen was completely clear of the box, complicating any interpretation of reaching behavior in delay conditions. Visual manipulations of cue salience could be tested in the presence and absence of delays in a computerized task (e.g., by presenting highly salient and less salient cues either in conjunction with a stimulus, or after a short delay), allowing researchers to avoid imposing physical obstacles that might alter normal reaching.

Additional investigations should also consider whether supplementary instructions and cue reminders continue to support inhibitory control in older children. Pupillometric estimations of mental effort suggest that young children (3.5 year olds) tend to reactively retrieve task relevant information in the moment, as needed, instead of maintaining information across time in preparation for future actions (Chatham, Frank, & Munakata, 2009). Thus, if 3-year-olds primarily employ a reactive form of control, they should benefit from cues (like tapping the cue, or placing it on the box lid) that trigger recall of task rules in the moments before they are required to respond. By contrast, older children who have transitioned to a more proactive form of control should be less likely to benefit from cues facilitating reactive retrieval (Chevalier & Blaye, 2009). Future work should investigate whether the developmental transition from primarily reactive to primarily proactive forms of control coincides with a decreased reliance on reminders highlighting the importance of cues 'in the moment', and increased reliance on task elements facilitating maintenance of relevant information across delays.

Finally, we leave open the question of whether the present findings will generalize to other response inhibition tasks. The present study considered inhibitory control in a relatively simple, two-rule (go, no-go) task. However, children may be more likely to rely on active computation in more difficult tasks, such as those requiring theory of mind (Carlson & Moses, 2001; Carlson, Moses & Claxton, 2004; Carlson, Moses, & Breton, 2002). Similarly, children may benefit from additional time in tasks requiring the resolution of multiple conflicting responses, or selection of a response from one of several active stimulus-response sets (e.g., Ridderinkhof & van der Molen, 1995; Kiselev, Espy & Sheffield, 2009; Tsujii & Watanabe, 2010). By examining the conditions under which children can be encouraged to inhibit prepotent responses, we may gain further insight into the processes underlying improvements in response inhibition across development.

Works Cited

Bates, D. M. (2007). Linear mixed model implementation in lme4. Manuscript, University of Wisconsin - Madison, January 2007.

Bates, D. M. (2013). Linear mixed model implementation in lme4. Manuscript, R-Project, September 2013. Accessed from ftp://www.r-project.org/pub/R/web/packages/lme4/vignettes/Implementation.pdf

Bates, D. M., & Sarkar, D. (2007). lme4: Linear mixed effects models using S4 classes. R package version 0.9975-12.

Berk, L.E. *(*1992*)*. Children's private speech: an overview of theory and the status of research. *In* R.M.Diaz *&* L.E.Berk *(Eds.), Private speech: From social interaction to self-regulation* (pp. 17–53*)*. Hillsdale, NJ: Erlbaum.

Blaye, A., & Chevalier, N. (2011). The role of goal representation in preschoolers' flexibility and inhibition. *Journal of Experimental Child Psychology*, 108, 469–483. doi: 10.1016/j.jecp.2010.09.006

Bluell, A. & Montgomery, D. E. (2013). The influence of stimulus discriminability on young children's interference control in the Stroop-like happy-sad task. Journal of Cognition & Development, published online February 25, 2013. doi: 10.1080/15248372.2013.767261

Bolker, B.M, Brooks, M. E., Clark, C. J., Geange, S. W., Poulsen, J. R., Stevens, M. H., & White, J. S. (2009). Generalized linear mixed models: a practical guide for ecology and evolution. *Trends in Ecology & Evolution*, 24, 127-135 doi: 10.1016/j.tree.2008.10.008.

Carlson S. M., Moses L. J. (2001). Individual differences in inhibitory control and children's theory of mind. *Child Development*, 72, 1032–1053. doi: 10.1111/1467-8624.00333

Carlson, S. M., Moses, L. J. and Breton, C. (2002). How specific is the relation between executive function and theory of mind? Contributions of inhibitory control and working memory. *Infant & Child Development*, 11, 73–92. doi: 10.1002/icd.298

Carlson, S. M., Moses, L. J. and Claxton, L. J. (2004). Individual differences in executive function and theory of mind: An investigation of inhibitory control and planning ability. *Journal of Experimental Child Psychology*, 87, 299–319. doi: 10.1016/j.jecp.2004.01.002

Chatham, C. H., Frank, M. J., & Munakata, Y. (2009). Pupillometric and behavioral markers of a developmental shift in the temporal dynamics of cognitive control. *Proceedings of the National Academy of Sciences,* 106, 5529–5533. doi: 10.1073/pnas.0810002106

Chevalier, N. & Blaye, A. (2009). Setting goals to switch between tasks: Effect of cue transparency on children's cognitive flexibility. *Developmental Psychology*, 45, 782–797. doi: 10.1037/a0015409

Chevalier, N., Martis, S. B., Curran, T., & Munakata, Y. (in prep). Proactive control can be encouraged during childhood: behavioral, ERP, and pupillometric evidence.

Davidson, M., Amso, D., Cruess Anderson, L. & Diamond, A. (2006). Development of cognitive control and executive functions from 4 to 13 years: evidence from manipulations of memory, inhibition, and task switching. *Neuropsychologia*, 44, 2037-2078. doi: 10.1016/j.neuropsychologia.2006.02.006

Diamond, A., Kirkham, N. & Amso, D. (2002). Conditions under which young children can hold two rules in mind and inhibit a prepotent response. *Developmental Psychology*, 38, 352–362.

Diamond, A. & Taylor, C. (1996). Development of an aspect of executive control: Development of the abilities to remember what I said and to "Do as I say, not as I do". *Developmental Psychobiology*, 29, 315–334.

Dowsett, S. M. and Livesey, D. J. (2000). The development of inhibitory control in preschool children: Effects of "executive skills" training. *Developmental Psychobiology*, 36, 161–174.

Duncan, J., Emslie, H., Williams, P., Johnson, R., & Freer, C. (1996). Intelligence and the frontal lobe: the organization of goal directed behavior, Cognitive Psychology, 30, 257–303.

Gerstadt, C. L., Hong, Y. J. & Diamond, A. (1994). The relationship between cognition and action: performance of 3.5 to 7 years old on Stroop-like day-night test. *Cognition*, 53, 129-153. doi: 10.1016/0010-0277(94)90068-X

Higa, W. R., Tharp, R. G., Calkins, R. P. (1978). Developmental verbal control of behavior: Implications for self-instructional training. Journal of Experimental Child Psychology, 26, 489–497. doi: 10.1016/0022-0965(78)90127-3

Jaeger, T. F. (2008). Categorical data analysis: Away from ANOVAs (transformation of not) and towards logit mixed models. *Journal of Memory & Language*, 59, 434–446.

Kiselev, S., Espy, K. A. & Sheffield, T. (2009). Age-related differences in reaction time task performance in young children. *Journal of Experimental Child Psychology*, 102, 150–166. doi: 10.1016/j.jecp.2008.02.002

Klenberg, L., Korkman, M., & Lahti-Nuuttila, P. (2001). Differential development of attention and executive functions in 3- to 12-year-old Finnish children. *Developmental Neuropsychology*, 20, 407–428. doi: 10.1207/S15326942DN2001_6

Kray, J., Eber, J. and Karbach, J. (2008). Verbal self-instructions in task switching: a compensatory tool for action-control deficits in childhood and old age? *Developmental Science*, 11: 223–236. doi: 10.1111/j.1467-7687.2008.00673.x

Laubrock, J., Engbert, R., Rolfs, M., and Kliegl, R. (2007). Microsaccades are an index of covert attention: commentary on Horowitz, Fine, Fencsik, Yurgenson, and Wolfe (2007). *Psychological Science* 18, 364–366. Discussion: 367–368. doi: 10.1111/j.1467-9280.2007.01904.x

Livesey, D.J., & Morgan, G.A. (1991). The development of response inhibition in 4- and 5-year-old children. *Australian Journal of Psychology*, 43, 133–137.

Luria, A. R. (1966). *Higher cortical functions in man*. New York: Basic.

Marcovitch, S., Boseovski, J. J. and Knapp, R. J. (2007), Use it or lose it: examining preschoolers' difficulty in maintaining and executing a goal. Developmental Science, 10: 559–564. doi: 10.1111/j.1467-7687.2007.00611.x

Marcovitch, S., Boseovski, J. J., Knapp, R. J. and Kane, M. J. (2010), Goal Neglect and Working Memory Capacity in 4- to 6-Year-Old Children. Child Development, 81: 1687–1695. doi: 10.1111/j.1467-8624.2010.01503.x

Martin-Rhee, M. M. & Bialystok, E. (2008). The development of two types of inhibitory control in monolingual and bilingual children. Bilingualism: Language & Cognition, 11, 81-93.

Meichenbaum, D. & Goodman, J. (1969). Reflection-Impulsivity and Verbal Control of Motor Behavior. *Child Development*, 40, 785-797.

Meichenbaum, D. H. & Goodman, J. (1971). Training impulsive children to talk to themselves: A means of developing self-control. Journal of Abnormal Psychology, 77, 115-126. doi: 10.1037/h0030773

Müller, U., Zelazo, P. D., Lurye, L. E., & Liebermann, D. P. (2008). The effect of labeling on preschool children's performance in the Dimensional Change Card Sort Task. Cognitive Development, 23, 395–408. doi: 10.1016/j.cogdev.2008.06.001

Morton, J. E. & Munakata, Y. (2002). Active versus latent representations: a neural network model of perseveration, dissociation, and decalage. Developmental Psychobiology, 40, 255-265. doi: 10.1002/dev.10033

Munakata, Y., Morton, J. B., & Yerys, B. E. (2003). Children's perseveration: attentional inertia and alternative accounts. Developmental Science, 6, 471–473.

Munakata, Y., Snyder, H. R., & Chatham, C.H. (2012). Developing cognitive control: Three key transitions. *Current Directions in Psychological Science,* 21: 71–77. doi: 10.1177/0963721412436807

Ridderinkhof, K. R. and van der Molen, M. W. (1995). A psychophysiological analysis of developmental Differences in the ability to resist interference. *Child Development*, 66, 1040–1056. doi: 10.1111/j.1467-8624.1995.tb00921.x

Simpson, A., & Riggs, K. J. (2011). Under what conditions do children have difficulty in inhibiting imitation? Evidence for the importance of planning specific responses. Journal of Experimental Child Psychology, 109, 512–524. doi: 10.1016/j.jecp.2011.02.015

Simpson, A., & Riggs, K. J. (2007). Under what conditions do young children have difficulty inhibiting manual actions? *Developmental Psychology,* 43, 417–428. doi: 10.1037/0012-1649.43.2.417

Simpson, A., & Riggs, K. J. (2006). Conditions under which children experience difficulty with a "button-press" go/no-go task. *Journal of Experimental Child Psychology*, 94, 18–26. doi: 10.1016/j.jecp.2005.10.003

Simpson, A., Riggs, K. J., Beck, S. R., Gorniak, S. L., Wu, Y., Abbott, D. & Diamond, A. (2012). Refining the understanding of inhibitory processes: how response prepotency is created and overcome. *Developmental Science*, 15, 62–73. doi: 10.1111/j.1467-7687.2011.01105.x

Towse, J. N., Lewis, C., & Knowles, M. (2007). When knowledge is not enough: The phenomenon of goal neglect in preschool children. *Journal of Experimental Child Psychology*, 96, 320–332. doi: 10.1016/j.jecp.2006.12.007

Tsujii, T. & Watanabe, S. (2010). Neural correlates of belief-bias reasoning under time pressure: A near infrared spectroscopy study. *Neuroimage*, 50, 1320–1326. doi: doi.org/10.1016/j.neuroimage.2010.01.026

Van der Ven, S. H. G., Kroesbergen, E. H., Boom, J., & Leseman, P. P. M. (2012). The development of executive functions and early mathematics: A dynamic relationship. *British Journal of Educational Psychology,* 82, 100–119. doi: 10.1111/j.2044-8279.2011.02035.x

Vygotsky, L. S. (1966, 1987). Thinking and speech. In *The collected works of L. S. Vygotsky: Vol. 1. Problems of general psychology* (N. Minick, Trans.). New York: Plenum Press.

Winsler, A., De León, J. R., Wallace, B. A., Willson-Quayle, A., Carlton, M. P. (2003). Private speech in preschool children: Developmental stability and change, across-task consistency, and relations with classroom behaviour. *Journal of Child Language*, 30, 583–608. doi: 10.1017/S0305000903005671

Appendix A: Summary Analyses of Task Performance

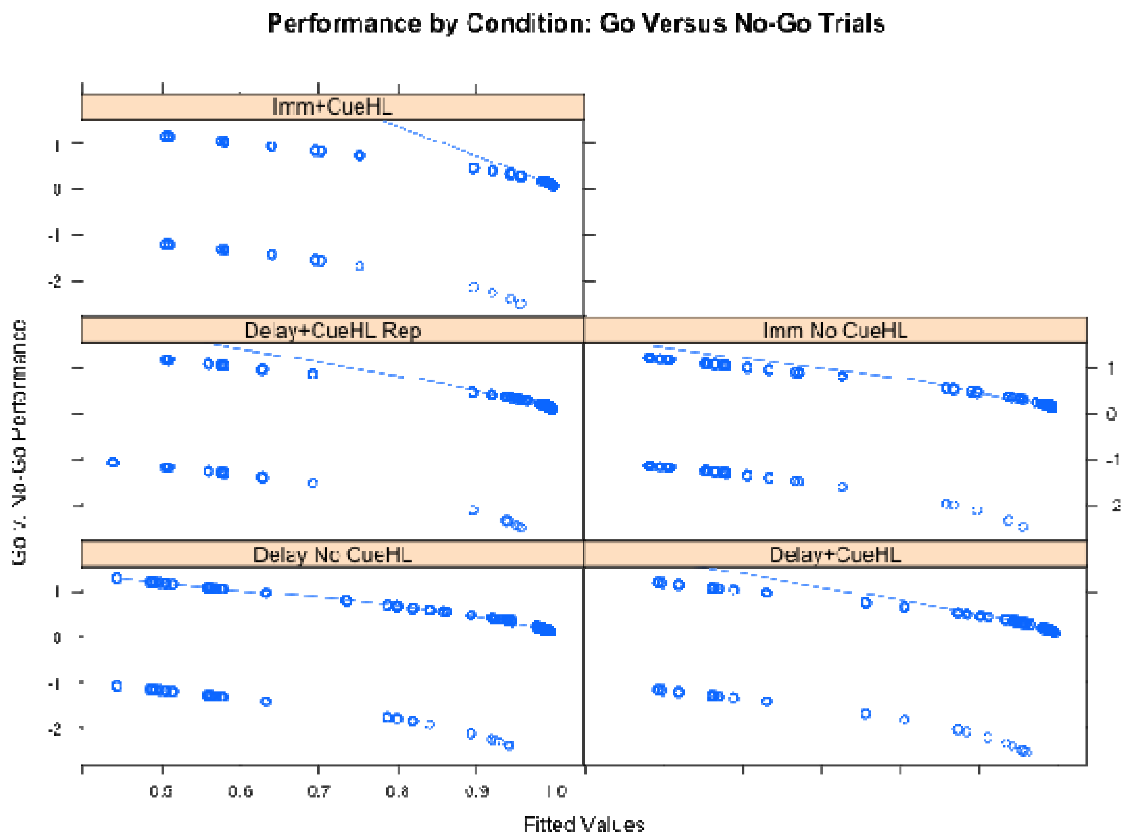

Figure A-1: Comparison of model fits by condition, where model is indicated by Performance ~ TrialType + Delay + CueHighlight+ Delay*TrialType + CueHighlight*TrialType + (1 | Subject), and fit is averaged across subjects and ages (see Model A-1 in Appendix A for parameter estimates). For each panel, the x-axis reflects log scaling; thus, each curve slopes down exponentially. Regression lines indicate fit by condition.

Table A-1: Summary of GLMM Fixed Effects, Basic Model

Predictor	Coefficient	*SE*	Wald *Z*	*p*
Intercept	5.94	(.421)	14.11	<.001***
Trial Type (Go = 1; No Go = -1)	2.56	(.174)	14.74	<.001***
Delay (Yes = 1; No = -1)	-0.19	(.391)	-0.50	.620
Cue Highlight (Yes = 1; No = -1)	1.01	(.392)	2.57	.010**
Delay * Cue Highlight	-0.21	(.394)	-0.53	.598
Age (Days), Mean Centered	0.011	(.004)	3.13	<.002**

Note: Intercept approximates average trial performance for a typical subject (i.e., a subject with average random effects). Model includes subject level random effects (not shown). N=150; Log-likelihood=-s449.91.

Table A-2: Summary of GLMM Fixed Effects, Basic Model (No-Go Intercept)

Predictor	Coefficient	*SE*	Wald *Z*	*p*
Intercept	3.43	(.419)	8.19	<.001***
Trial Type (Go = 1; No Go = 0)	5.97	(.621)	9.61	<.001***
Delay (Yes = 1; No = -1)	-0.04	(.418)	-0.10	.924
Cue Highlight (Yes = 1; No = -1)	1.11	(.408)	2.73	<.007**
Delay * Trial Type	-1.67	(.619)	-2.70	<.008**
Cue Highlight * Trial Type	-0.63	(.371)	-1.70	<.09
Age (Days), Mean Centered	0.012	(.004)	3.07	<.003**

Note: Intercept approximates average no-go trial performance for a typical subject (i.e., a subject with average random effects). Model includes subject level random effects (not shown). N=150; Log-likelihood=-438.56.

Table A-3: Summary of GLMM Fixed Effects, Model A-1 + Attention Score

Predictor	Coefficient	*SE*	Wald *Z*	*p*
Intercept	-2.11	(2.34)	-0.902	.367
Trial Type (Go = 1; No Go = 0)	5.67	(.449)	12.64	.140
Delay (Yes = 1; No = -1)	2.69	(2.47)	1.09	.277
Cue Highlight (Yes = 1; No = -1)	.0779	(2.29)	0.34	.734†
Total Attention Score	0.383	(.164)	2.33	<.02*
Age (Days), Mean Centered	0.009	(.004)	2.49	<.02*
Attention x Delay	-0.191	(.173)	-1.10	.270
Attention x Cue Highlight	0.021	(.164)	0.13	.901

Note: Intercept approximates average trial performance for a typical subject (i.e., a subject with average random effects). Model includes subject level random effects (not shown). N=135; Log-likelihood=-367.93. † : Cue Highlight was made insignificant because of poor model fit; when interactions are excluded (Attention*CueHL, Attention*Delay), Cue Highlight $p < .02$; Delay $p > .9$; Attention $p < .052$.

Table A-4: Summary of GLMM Fixed Effects, Model A-1 + Experimental Setting

Predictor	Coefficient	*SE*	Wald *Z*	*p*
Intercept	3.32	(.399)	8.34	<.001***
Trial Type (Go = 1; No Go = 0)	5.12	(.351)	14.62	<.001***
Delay (Yes = 1; No = -1)	-0.15	(.405)	-0.36	.720
Cue Highlight (Yes = 1; No= -1)	0.67	(.396)	0.396	.09†
Setting (Museum = 1; Lab = -1)	-1.71	(.688)	-2.49	<.02*
Age (Days), Mean Centered	0.009	(.003)	2.96	<.004**
Setting x Delay	0.128	(.694)	0.18	.854
Setting x Cue Highlight	0.642	(.668)	0.96	.337

Note: Intercept approximates average no-go trial performance for a typical subject (i.e., a subject with average random effects). Model includes subject level random effects (not shown). N=150; Log-likelihood=-448.43.
† : Cue Highlight was made insignificant because of poor model fit; when interactions are excluded (Setting*CueHL, Setting*Delay), Cue Highlight $p < .01$; Delay $p > .6$; Setting $p < .052$.

Table A-5: Summary of GLMM Fixed Effects, Spoken Language Model

Predictor	Coefficient	*SE*	Wald *Z*	*p*
Intercept	5.73	(.469)	12.23	<.001***
Trial Type (Go = 1; No Go = 0)	2.80	(.219)	12.79	<.001**
Delay (Yes = 1; No = -1)	-0.26	(.394)	-0.66	.507
Cue Highlight (Yes = 1; No = -1)	1.12	(.385)	2.91	<.004**
Age (Days), Mean Centered	0.012	(.004)	2.91	<.004**
Total Task Relevant Verbal Utterances Score	0.215	(.112)	1.93	.054

Note: N=135. Note: Intercept approximates average trial performance for a typical subject (i.e., a subject with average random effects). Model includes subject level random effects (not shown). Log-likelihood=-371.91.

Table A-6: Summary of Linear Model Predictions, Go Trial Reaction Times

Predictor	Coefficient	*SE*	t value	*p*
Intercept	1.474	(.583)	2.53	.013*
Total Attention Score	-0.061	(.020)	-3.08	.003**
Total No-Go Trials Correct	0.037	(.016)	2.40	.018*
Age (Days), Mean Centered	-0.0007	(.000)	-1.55	.125
Cue Highlight (Yes=1; No=-1)	-0.211	(.047)	-4.46	<.001***
Delay (Yes=1; No = -1)	0.187	(.047)	3.97	<.001***

Note: N=135. Three outliers replaced with overall mean (see Figure B-1 in Appendix B for details). Adjusted $R^2 = .27$.

Appendix B: Go-Trial Reaction Time Outlier Analysis

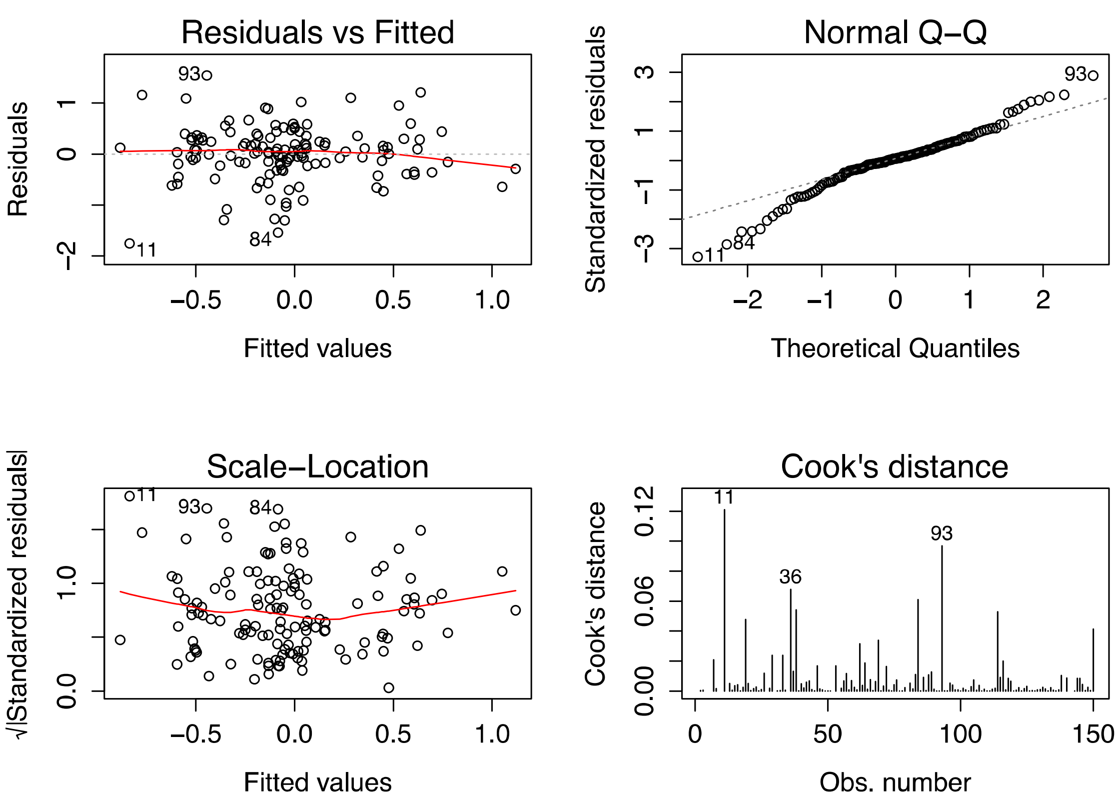

Figure B-1: Examination of go-trial outliers in full reaction time model, where model = log(MeanGoRT ~ Attention Score + No Go Correct (count) + Delay + CueHL + AgeDays. From left to right, moving counterclockwise: in the first two figures (Residuals versus Fitted and Scale-Location plots), errors appear normally distributed (i.e., they do not tend to increase or decrease with the size of predicted values). Three outliers are identified (observations 11, 84, and 93). Estimates of Cook's Distance, which factors in the contributions of each observation's studentized deleted residual value (an assessment of whether an observation is unusual with respect to a model determined by all of the remaining observations) and lever value (the weight an observation has in determining the overall model) to overall model fit, suggest that observations 11, 36, and 93 represent outliers. Finally, the Q-Q plot indicates that three outliers (11, 84, and 93) increase the thickness of the model tails. Thus, three observations (11, 84, and 93) were consistently identified as outliers, and replaced with the overall go trial mean (1.175 s) in the analysis.

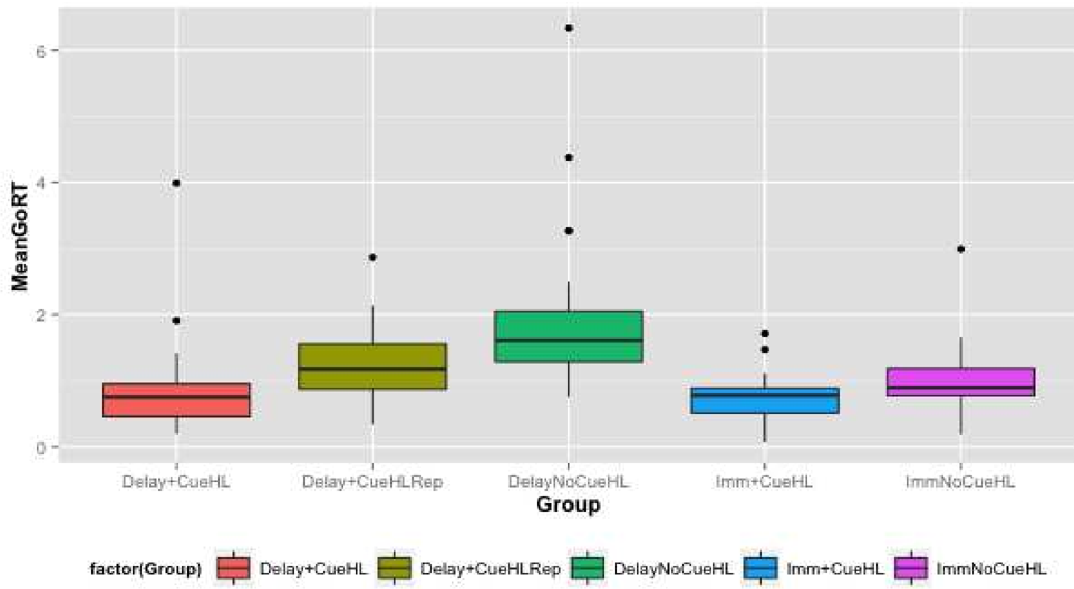

Figure B-2: Average *go* trial reaction times by condition (including outliers). Children in Cue Highlight conditions reached more quickly, on average, than children in No Cue Highlight conditions, and children in Delay conditions reached more quickly than children in Immediate conditions.