The Timescale of Control: A Meta-Control Property that Generalizes across Tasks but Varies between Types of Control

Abhishek Dey [1] · Julie M. Bugg [1]

Abstract
Prominent models of control assume that conflict and the probability of conflict are signals used by control processes that regulate attention. For example, when conflict is frequent across preceding trials (i.e., high probability of conflict), control processes bias attention toward goal-relevant information on subsequent trials. An important but underspecified question regards the meta-control property of timescale—that is, how far back does the control system "look" to determine the probability of conflict? To address this question, Aben, Verguts, and Van den Bussche (2017) developed a statistical model quantifying the timescale of control. In a flanker task, they observed short timescales for lists with a low probability of conflict (which induce reactive control) and long timescales for lists with a high probability of conflict (which induce proactive control). To investigate the domain generality of these timescales, we applied their model to two additional conflict tasks that manipulated the list-wide probability of conflict. Our findings replicated Aben et al. suggesting meta-control may be task general with respect to timescales operating on the list level. We subsequently modified their model to examine timescale differences for *items* in the same list that differed in their probability of conflict but not the type of control engaged. We failed to detect a difference in timescales between items. Collectively, the findings demonstrate that differences in the timescale of control are task general and suggest that timescale differences are driven by the type of control engaged and not by the probability of conflict per se.

Keywords Cognitive control · Timescale of control · Learning rate · List-wide proportion congruence · Item-specific proportion congruence

Cognitive control is the ability to pursue goal-directed behavior in the face of more habitual or immediately compelling alternative behaviors (Cohen, 2017). Several prominent models of cognitive control assume that conflict serves as an important signal for adjustments in control (Blais et al., 2007; Botvinick et al., 2006; Verguts & Notebaert, 2008). Conflict occurs when relevant (to-be-attended) and irrelevant (to-be-

ignored) stimulus dimensions activate competing responses, such as when naming the color on an incongruent trial (e.g., RED in blue ink) in the Stroop task or when having to bypass one's usual exit that leads home in order to fulfill the goal of picking up milk at the grocery store.

The term *adaptive control* implies that the control system learns about the probability of conflict within a given context and modulates attention accordingly (Braem et al., 2019; for recent reviews, see Bugg & Crump, 2012; Bugg, 2017). Indeed, many extant models view the probability of conflict, and thus past experiences with conflict, as central to how the system modulates attention. However, what remains underspecified is the question of how far back the control system "looks" to learn about the probability of conflict. Specifically, what is the timescale under which the control system operates? The timescale of control can be thought of as the window size representing how many prior trials are accumulated to calculate the probability of conflict for an upcoming trial. We consider timescale to be a key meta-control property of the cognitive control system, and we aim to further characterize this property by addressing three questions using

✉ Abhishek Dey
  dey.a@wustl.edu

[1] Department of Psychological and Brain Sciences, Washington University in St. Louis, Campus Box 1125, St. Louis, MO 63130, USA

data from confound-minimized experiments (Braem et al., 2019): 1) Does the timescale of cognitive control vary across lists that differ in their probabilities of conflict, as a recent model proposed and initial data supported (Aben et al., 2017), and why might it vary? 2) Is the effect of the list-wide probability of conflict on the timescale of control consistent across different conflict tasks? 3) Are differing timescales of control also observed for items within a list that vary in their probability of conflict? The present study represents the most comprehensive investigation of this meta-control property to date and aims to provide new information about the generality and specificity of meta-control.

Timescale of Control: A Meta-control Property

For the most part, the timescale of control has been assumed to be a fixed value. For example, the conflict-monitoring model (Botvinick et al., 2001) and the item-specific adaption of the conflict-monitoring model (Blais et al., 2007) both use a fixed λ weighting parameter to represent how much experiences with previous conflict adjust control for an upcoming trial or item. This fixed λ parameter essentially captures the timescale of control such that a small λ value assumes a small window size or short timescale. Conversely, a large λ value assumes a large window size or long timescale.

However, assuming a fixed timescale for participants in all experimental conditions yields sub-optimal estimates of behavior (Jiang et al., 2014). Jiang et al. (2014) developed a Bayesian model of flexible control, which takes inspiration from reinforcement learning models, and assumed that the probability of conflict is calculated using a learning rate parameter (α). Conceptually, this learning rate parameter can be thought of as the complement of λ (i.e., $1 - \lambda$) in that it weights how much the presence of conflict on the current trial adjusts control for an upcoming trial. In other words, a large learning rate value indexes a short timescale, and a small learning rate value indexes a long timescale.

Notably, variation (across individuals and trials) in parameters representing the timescale of control also has been shown to correlate with neural activity. For example, Jiang et al. (2015) applied the Bayesian model of flexible control to extract a timescale parameter that varied trial-by-trial for each participant performing a face-word Stroop task. In their task, participants were asked to respond to the gender of a face while ignoring the word "male" or "female" superimposed on the face. Using whole-brain searchlight techniques, they found that variation in timescales correlated with left anterior insula and inferior frontal gyrus. They concluded that these regions are integral for determining timescales necessary for predicting conflict which is consequently used to modulate control. This suggests that the timescale of control is

not merely useful for mathematical models of control but may also be a meta-control property that is represented neurally.

Assuming the timescale of control is flexible, a question that has not received much attention is does it vary systematically between situations that evoke different control states? In other words, do timescales not only vary between individuals and within a run of trials (as alluded to earlier), but also between experimental conditions that induce different *types* of control?

Empirically Determining the Timescale of Control for Different Types of Control

The dual-mechanisms of control account (Braver et al., 2007; Braver, 2012) states that there are two distinguishable types of control: reactive control and proactive control. According to Braver (2012) "reactive control reflects transient stimulus-driven goal reactivation…based on interference demands or episodic associations. Proactive control reflects the sustained and anticipatory maintenance of goal-relevant information… to enable optimal cognitive performance." (p. 106). In a reactive state, how much a person is paying attention to the target dimension can vary dramatically trial-by-trial. By contrast, in a proactive state, the assumption is that a person maintains a stable level of attention to the target dimension across trials.

An experimental method by which one can induce different control types is to parametrically change the global probability of conflict within a list. For example, one can create lists that are either mostly congruent (low probability of conflict) or mostly incongruent (high probability of conflict). A robust finding is that performance differs across lists such that congruency effects (e.g., difference between incongruent and congruent reaction times in tasks, such as Stroop, flanker, and Simon) are significantly increased in lists with a low (i.e., mostly congruent) relative to high (i.e., mostly incongruent) probability of conflict (Kane & Engle, 2003; Lindsay & Jacoby, 1994; Logan & Zbrodoff, 1979; Logan et al., 1984; Lowe & Mitterer, 1982; Toth et al., 1995; for reviews see Bugg, 2012; Bugg & Crump, 2012; Bugg, 2017). Critically, this *list-wide proportion congruence effect* holds for novel, frequency-matched sets of items that are 50% congruent in each list (i.e., diagnostic items; Bugg, 2014; Bugg & Chanani, 2011; Hutchison, 2011; Gonthier, Braver, & Bugg, 2016; cf. Bugg et al., 2011b), which control for a number of confounds that could otherwise potentially explain the effect (e.g., item-specific contingency learning; item-specific control; bottom-up priming of control states; for discussion see Braem et al., 2019). The prevailing interpretation of this list-wide proportion congruence effect is that the control system implicitly learns about the statistical regularities of conflict within a list and biases attention based on those regularities (Blais, Harris, Guerrero, & Bunge, 2012). In mostly congruent

lists, where conflict is not anticipated, control is low, but the occurrence of incongruent trials produces interference that transiently (reactively) activates the task goal. In contrast, in mostly incongruent lists, there is a sustained (proactive) heightening of control (i.e., decreased attention to the irrelevant dimension and/or increased attention to the relevant dimension) in anticipation of conflict.

Supporting this characterization, in an fMRI study, De Pisapia and Braver (2006) found that lists with a low probability of conflict showed behavioral and neural markers of reactive control, whereas lists with a high probability of conflict showed markers of proactive control. Specifically, they found increased *transient* activity in the prefrontal cortex on incongruent trials when participants performed a color-word Stroop task with a low probability of conflict at the list-level. In juxtaposition, they found increased *sustained* activity in the prefrontal cortex when participants performed the task with a high probability of conflict at the list-level. Simulation results from their computational model, which included two conflict detection units—one that calculated conflict on a short timescale and one that calculated conflict on a longer timescale—fit the data well. However, their model did not quantify just how far back the control system "looked" to learn about the probability of conflict in the different list types.

In line with the research of De Pisapia and Braver (De Pisapia & Braver, 2006), Aben et al. (2017) posited that conditions that induce transient (reactive) control (e.g., lists with a low probability of conflict) should operate on a short timescale such that conflict in a small window of prior trials affects control on the current trial; in contrast, conditions that induce sustained (proactive) control (e.g., lists with a high probability of conflict) should operate on a long timescale such that a larger window of prior trials affects control on the current trial (i.e., "protracted occurrences of conflict [engage] sustained proactive control"; De Pisapia & Braver, 2006, p. 1326). Critically, they put these predictions to the test by developing a statistical model to empirically quantify the timescale of control in lists that differed in their probability of conflict. We refer to the model that they developed as the extended congruency sequence effect (CSE) model.

The extended CSE model takes its name from the CSE— the observation that the congruency effect (i.e., the difference in reaction time between congruent and incongruent trials) in conflict tasks is subject to cross-trial sequence effects. The conflict status of the previous trial ($C_{i-1,}$) influences the degree to which conflict on the current trial (C_i) affects RT and accuracy (Grattonet al., 1992; for reviews, see Duthoo et al., 2014; Egner, 2007). In short, the CSE is the impact of C_{i-1} on C_i (i.e., the C_iC_{i-1} interaction). Figure 1 depicts the CSE pattern whereby the congruency effect is reduced when the previous trial is incongruent compared with when the previous trial is congruent. One prominent interpretation of the CSE is that it reflects local adjustments of attention by the control system
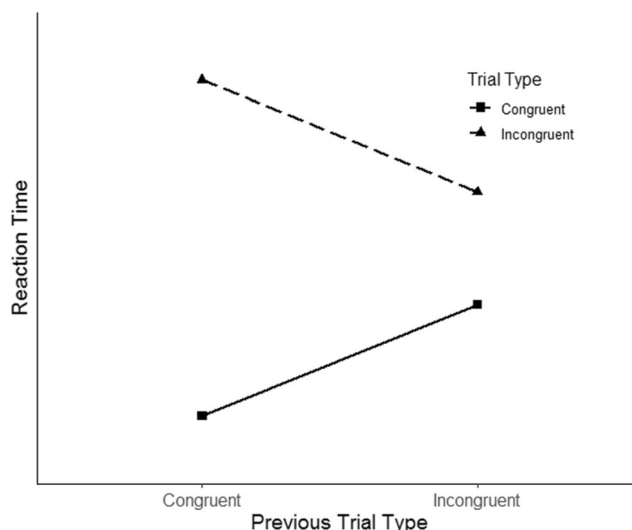

Fig. 1 Hypothetical data depicting the congruency sequence effect (CSE). Plotted on the x axis is the trial type of the preceding trial. Squares depict current congruent trials. Triangles depict current incongruent trials. The difference between squares and triangles for a given previous trial type depicts the congruency effect. This plot shows that the congruency effect is reduced when the previous trial is incongruent compared to when the previous trial is congruent.

based on the detection of conflict (Botvinick et al., 2001; see also Egner, 2007 for review). That is, based on the conflict status of the previous trial (trial i-1), the system adjusts attention to the current trial (trial i). When encountering conflict on trial i-1, the control system increases attention to the relevant stimulus dimension, thereby reducing the congruency effect for the current trial.

The unique contribution of the model of Aben et al. (2017) was that it considered not just the impact of the congruency state of the first previous trial on the current trial but also the independent effect of multiple previous trials on the congruency effect on the current trial (Colvett et al., 2020; Durston, Davidson, Thomas, Worden, Tottenham, Martinez, Watts, Ulug, & Casey, 2003; Horga, Maia, Wang, Wang, Marsh, & Peterson, 2011; Jiménez & Méndez, 2013; Jiménez & Méndez, 2014). In other words, they looked at the independent effect of C_{i-1}, C_{i-2}, C_{i-3}, … C_{i-k} on C_i. In their model, k was equal to the maximum trial distance that had a significant impact on the current trial. The primary parameters of interest were current trial by previous trial interactions (e.g., C_iC_{i-1}, C_iC_{i-2}, … C_iC_{i-k}), and the value of an interaction term for a given trial distance, which was deemed to be that trial distance's *conflict-adaptation-weight* (CAW).

In the extended CSE model, the CAW for a given trial distance quantifies that trial distance's *independent* impact on the current trial's congruency effect. Critically, the relative change in CAWs across trial distances (slopes in Fig. 2) was interpreted by Aben et al. (2017) as an index of the timescale of control. A short timescale was indexed by CAWs that changed dramatically over trial distance, whereas a long timescale was
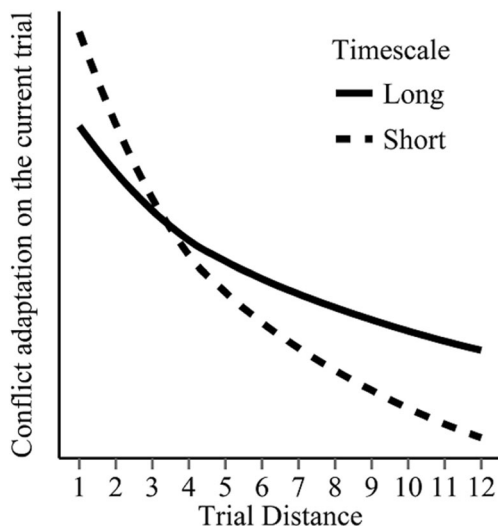
Fig. 2 Hypothetical influence of previous trials' congruency state on the current trial's congruency effect. Conflict adaptation on the current trial, otherwise known as the conflict-adaptation weight (CAW) is plotted on the y-axis. Trial distance is plotted on the x-axis. The relative changes in CAW across trial distance (i.e., the slopes of the trajectories) index the timescale of control. Long timescales are indexed by relatively smaller changes across trial distance. Short timescales are indexed by relatively steeper changes across trial distance. Source: Aben et al., 2017.

indexed by CAWs that did not change much over trial distance. Conceptually, a short timescale translates to a relative increase in the weighting of recent trials and a relative decrease in the weighting of distal trials when compared to a long timescale. Figure 2 (from Aben et al., 2017) illustrates the difference between short and long timescales as captured by their model.

The key finding of Aben et al. (2017) was that the probability of conflict within a list influenced the timescale of control. Lists with a low probability of conflict (i.e., flanker lists with a proportion congruence of 0.80 [PC 80]) induced a shorter timescale of control relative to lists with a high probability of conflict (i.e., PC 20 flanker lists). This suggests that adaptation to conflict occurs on different timescales with relatively greater weighting of recent trials when conflict is improbable and relatively greater weighting of distal trials when conflict is probable. As Aben et al. noted, these findings are compatible with theories that distinguish transient and sustained control, respectively, such as the dual-mechanisms of control account.

Present Research

The results of Aben et al. (2017) represented the first piece of evidence that the timescale of control may vary across lists that differ in their probability of conflict and consequently, in the case of the list-wide PC manipulation, the type of control engaged (i.e., reactive vs. proactive). However, to date, these findings have not yet been replicated. Thus, it remains

uncertain whether their findings are stable or specific to the task and design they employed. Aben et al. investigated timescales using a two-choice flanker task that included only biased (inducer) items (e.g., flanker stimuli with left flanking arrows were mostly congruent in the mostly congruent lists and mostly incongruent in mostly incongruent lists). Such designs leave open the possibility that processes such as contingency learning (i.e., processes by which participants predict responses [central arrow identity] based on flanking arrows, for example, pressing right key whenever left flanking arrows are encountered in mostly incongruent lists; see Schmidt & Besner, 2008), may explain the difference in the congruency effect across lists (Braem et al., 2019; Bugg, 2012) and potentially the differing timescales. By including diagnostic items that are frequency and PC matched across lists (e.g., 50% congruent in each list), one can confirm that performance differences between lists are indicative of the expected differences in control (e.g., proactive control in the mostly incongruent list).

Motivated by the findings of Aben et al. (2017), in the present research we attempted to replicate the timescale patterns they observed in experimental conditions that demonstrably pushed participants to engage in either reactive or proactive control (as confirmed by performance on the diagnostic items). In Analyses 1 and 2, we sought to determine whether the list-level timescale patterns observed by Aben et al. are replicable and task-general. We did this by applying their statistical model to data from two conflict tasks that differed from the flanker task they used: a picture-word Stroop task and a color-word Stroop task. Based on the findings of Aben et al., in Analyses 1 and 2, it was predicted that the timescale of control would be shorter in conditions that induce transient/reactive control (mostly congruent lists) and longer in conditions that induce sustained/proactive control (mostly incongruent lists).

Finally, in Analysis 3, we examined the novel question of whether differences in the probability of conflict between items (rather than between lists) lead to differences in timescales of control. We accomplished this by examining timescales of control within a 50% congruent list of picture-word Stroop trials in which item-specific proportion congruence was manipulated. Briefly, certain pictures were mostly congruent within that list and other pictures were mostly incongruent, and they were randomly intermixed. The *item-specific proportion congruence* effect is the reduction in the congruency effect for mostly incongruent compared to mostly congruent items (Jacoby et al., 2003). Critically, this effect reflects reactive control based on retrieval of learned control settings (i.e., episodic associations) that are specific to each item (e.g., mostly incongruent items retrieve a setting entailing less processing of irrelevant dimension and/or more processing of target). Investigating this question allowed us to assess the generalizability of timescales, particularly whether different timescales of control would be observed for mostly congruent and mostly incongruent items

like the differences found for mostly congruent and mostly incongruent lists.

Analysis 1: Examining the List-Wide Timescale of Control in a Picture-Word Stroop Task from Gonthier et al. (2016)

For Analysis 1, we used data from a picture-word Stroop task in which list-wide proportion congruence was manipulated and participants named an animal in a picture while ignoring a superimposed animal word (Gonthier et al., 2016).[1] The Gonthier et al. study included a large sample ($N = 93$), within-subjects manipulations of the factors of interest, and diagnostic items, which as detailed earlier are important for confirming that differences between lists are attributable to differences in control (Braem et al., 2019). Specifically, the data showed that the diagnostic items within the mostly incongruent list had smaller Stroop effects than the diagnostic items within the mostly congruent list, consistent with the view that participants adopted proactive control in the mostly incongruent list (De Pisapia & Braver, 2006; Gonthier et al., 2016). We applied the extended CSE model and investigated whether the timescale of control varied systematically between mostly incongruent and mostly congruent lists in the picture-word Stroop task. A secondary goal was to compare a list-wide PC 50 list to the mostly incongruent and mostly congruent lists.

Method

Each participant was exposed to three conditions (for a more detailed description of the methods, see Gonthier et al., 2016). The conditions comprised a list-wide mostly congruent condition, a list-wide mostly incongruent condition, and a list-wide PC 50 condition. All three conditions included diagnostic items (four animal pictures that were PC 50; e.g., cow, frog, pig, seal).[2] What differed across conditions was the PC of the other four animal pictures (i.e., the inducer items; e.g., bird,

cat, dog, fish) included in each list. In the list-wide mostly congruent condition, the other four were PC 75. When combined with the diagnostic items, this led to an overall PC 67 list. In the list-wide mostly incongruent condition, the other four were PC 25. This created an overall PC 33 list. The list-wide mostly congruent and mostly incongruent conditions were comprised of 384 trials. For the third condition, the list-wide PC 50 condition, the other four pictures were two PC 75 items and two PC 25 items. The latter two sets of items were used to manipulate item-specific proportion congruence within this list. For the purposes of Analysis 1, which focused on timescales associated with the list-wide probability of conflict, this third condition was treated as a PC 50 list for comparison to the PC-67 and PC-33 lists. The list-wide PC 50 condition was comprised of 432 trials.

Of the original 93 participants included in this dataset 83 were retained for analysis. 10 participants were excluded, because they did not complete all the conditions within the experimental paradigm.

Analytical approach

Unless otherwise mentioned, the analytical approach mirrored the approach detailed in Aben et al. (2017). The analysis included all items (inducer and diagnostic).[3] The following trials were excluded: the first trial of each condition (0.41%), error trials (3.77%), and trials following errors (3.77%). Like Aben et al., we also box-trimmed the RT data; however, we used the cutoffs that we have used previously in our research on the Stroop task (Bugg, 2014; Bugg & Dey, 2018; Gonthier et al., 2016). Trials <200 ms and >3,000 ms were excluded (0.34%).

Hierarchical linear models (HLMs) were used to analyze RTs. These models require an assumption that the residuals of the dependent variable are normally distributed, but RT residuals are not typically normally distributed. Thus, we inverse transformed RTs (1/RT) and multiplied them by $-10,000$ to restrict the number of decimal places (Kinoshita, Mozer, & Forster, 2011). Smaller inverse RTs reflect faster responses, so the direction of the transformed scale is consistent with raw RTs.[4]

---

[1] For this and all subsequent analyses, we selected data sets based on three major criteria. 1) They had to be highly powered with $Ns > 60$. 2) They were collected using fully within-subject manipulations of the factors of interest so that we could appropriately model timescales using the approach specified in Aben et al. (2017). 3) They were collected using designs that minimized known confounds and included diagnostic items that enabled inferences about control (see Braem et al., 2019 for review). Although there are three other published datasets that met the latter criterion (Bugg, 2014; Bugg & Chanani, 2011; Hutchison, 2011), none additionally met the remaining two criteria.

[2] Note that behavioral patterns in Gonthier et al. (2016) indicated list-level control. Specifically, the same PC 50 items in a list-wide mostly congruent block had a larger Stroop effect (115 ms) than when in a list-wide mostly incongruent block (86 ms). This pattern for diagnostic items has been taken to indicate the operation of a list-level control mechanism as opposed to other mechanisms (Bugg, 2014; Braem et al., 2019; e.g., contingency learning; see also Cohen-Shikora, Suh, & Bugg, 2019, for evidence countering a temporal learning mechanism).

[3] There were too few trials to restrict the analysis to diagnostic items only in the current and subsequent analyses.

[4] Because it has been shown that using nonlinear transformations, such as inverse transformations, can systematically alter interaction effects compared with using untransformed data (Balota et al., 2013), Lo and Andrews (2015) suggested using generalized linear mixed-effect models (GLMMs) with a Gamma distribution link function as an alternative to nonlinear transformations. For the primary analyses in the present research (i.e., those reporting the timescale of control), analyses of transformed data are reported for purposes of examining whether our results replicated Aben et al. (2017). However, we also modified the model using untransformed data with a Gamma link GLMM. For the interested reader, following recommendations in the literature (Balota et al., 2013), the results from the modified model are reported in our supplementary materials.

We first sought to determine how many previous trials back ought to be included to model the timescale of control. We ran an HLM across all conditions which included congruency of the current trial (C_i), congruency of the kth previous trials (C_{i-k}), and the interactions of the current trial and k^{th} previous trials ($C_i C_{i-k}$). To be conservative, we chose to look at up to 24 previous trials back as an initial starting point. Because of model convergence issues, the HLM was implemented with only the intercept being allowed to vary on a subject level. This resulted in the level one equation:

$$RT = \beta_0 + \beta_1 C_i + \beta_2 C_{(i-1)} + \beta_3 C_{(i-2)} \ldots + \beta_{25} C_{(i-24)}$$
$$+ \beta_{26} C_i C_{(i-1)} + \beta_{27} C_i C_{(i-2)} \ldots + \beta_{49} C_i C_{(i-24)} \qquad (1)$$

The results of this analysis are shown in Fig. 3. Based on this analysis we chose to include up to 16 trials back given the t-values were around 1.96 until trial 16 (with a trial distance of 10 and 20 being somewhat substantial exceptions). Following this preliminary analysis, RTs were modeled at two hierarchically related levels. The level one equation of the model was identical to that of Equation 1 with the caveat that k was reset from 24 to 16. This yielded the level one equation:

$$RT = \beta_0 + \beta_1 C_i + \beta_2 C_{(i-1)} + \beta_3 C_{(i-2)} \ldots + \beta_{17} C_{(i-16)}$$
$$+ \beta_{18} C_i C_{(i-1)} + \beta_{19} C_i C_{(i-2)} \ldots + \beta_{33} C_i C_{(i-16)} \qquad (2)$$

Equation 2 was applied iteratively to each subject within each condition (list-wide mostly congruent condition, list-wide PC 50 condition, list-wide mostly incongruent condition). In doing so, the coefficients of the interaction terms for each subject within each condition were extracted. These coefficients are the conflict adaptation weights (CAWs) discussed earlier. Each CAW at a given trial distance reflects the magnitude of shift in the congruency effect on the present trial if an incongruent trial was presented at that trial distance. At the second level, the 16 CAWs estimated by Eq. 2 were entered in as dependent variables in an HLM with trial distance and condition as predictors. Trial distance was log transformed and then subsequently mean centered to allow for better interpretation of the intercept in the model.[5]

The effect of interest is the interaction between trial distance and condition. To assess the significance of this effect, predictors in the second level were entered in a stepwise fashion and each model was tested against its previous simpler nested model. Given the nested nature of the models, a test

---

[5] A fully random HLM structure could not be implemented without encountering convergence issues. That is, trial distance, condition, and the trial distance by condition interaction could not be entered in as random effects. This was also the case in Aben et al. (2017). The effect of condition could be allowed to be random, but we chose only to allow the intercept to be random to more closely follow the procedure in Aben et al. Allowing condition to be random or not did not significantly change the results of this analysis.
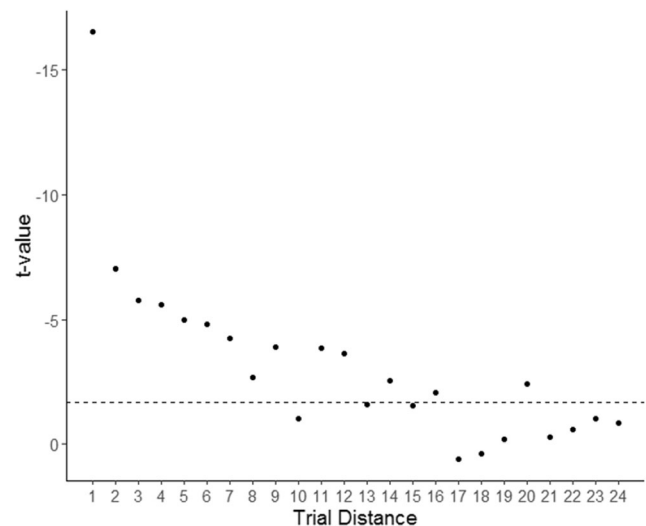

Fig. 3 $C_i C_{i-k}$ interaction t values from Eq. 1 plotted as a function of trial distance (ranges from 1-24). The dotted line represents the critical two-tailed t-value of 1.96 given infinite degrees of freedom. Data are from list-wide manipulations of a picture-word Stroop task from Gonthier et al. (2016).

of the log likelihood ratio determined if the models were significantly different from one another with larger log likelihoods indicating more variance explained. Nonetheless, the Akaike information criterion (Akaike, 1974) also is reported as a measure of model fit to mimic Aben et al. (2017). A smaller Akaike information criterion (AIC) indicates a better fit.

Results

The summary of the model comparisons is shown in Table 1. By testing Model 1 versus Model 0, we found that CAWs significantly differed as a function of log trial distance, $X^2(1) = 99.19$, $p < 0.001$. The test of Model 2 versus Model 0 indicated that CAWs also significantly differed as a function of list condition, $X^2(2) = 22.46$, $p < 0.001$. The model with both log trial distance and list condition explained the data better than the model with just log trial distance, $X^2(2) = 23.04$, $p < 0.001$ (Model 3 vs. Model 1), and the model with just list condition, $X^2(1) = 99.77$, $p < 0.001$ (Model 3 vs. Model 2). Critically, there was a trending effect when comparing the full model, which included the interaction of log trial distance and list condition (Model 4), and the main effects only model (Model 3), $X^2(2) = 5.40$, $p = 0.067$. This warranted further analysis as to how the interaction terms (or slopes of the list conditions over trial distance) differed.

The regression coefficients of the full model are displayed in Table 2. Figure 4 displays the estimates of the full model corrected for the intercept of each condition (following Aben et al., 2017). We ran a no-intercept model and used the coefficients (mean slope for each condition) in a linear combination to determine if the slopes were different from one another.

Table 1 Gonthier et al. (2016) List-Wide Proportion Congruence Model Comparisons with Inverse Transformed RT

Model	df	AIC	Log lik.	Test	X^2	p
0. Intercept	3	8077	-4036			
1. Log trial distance	4	7980	-3986	1 vs. 0	99.19	<0.001
2. Condition	5	8058	-4024	2 vs. 0	22.46	<0.001
3. Log trial distance + Condition	6	7961	-3975	3 vs. 1	23.04	<0.001
				3 vs. 2	99.77	<0.001
4. Log trial distance + Condition + Log trial distance x Condition	8	7960	-3971	4 vs. 3	5.40	0.067

AIC = Akaike Information criterion; log lik. = log likelihood.

This was accomplished using the glht() function from the multcomp package in R (Hothorn, Bretz, & Westfall, 2008). Because degrees of freedom are difficult to estimate, the analysis uses z score estimates to determine significant differences. The p values were adjusted using the Holm's procedure.

The list-wide mostly congruent slope was significantly steeper than the list-wide mostly incongruent slope, z = 2.98, p = 0.009. Thus, we observed a shorter timescale of control in the list-wide mostly congruent condition relative to the list-wide mostly incongruent condition. The list-wide mostly congruent slope and the list-wide PC 50 slope were not significantly different, z = 0.97, p = 0.332, indicating no differences in the timescale. The list-wide PC 50 slope differed marginally from the list-wide mostly incongruent slope, z = 2.01, p = 0.090.

Discussion

The results from Analysis 1 are generally consistent with the findings from Aben et al. (2017). A shorter timescale of control (steeper slope) was observed for the list-wide mostly congruent condition and a longer timescale of control was observed for the list-wide mostly incongruent condition. These results indicate that the meta-control system places more importance or weight on recent events during low-probability conflict situations when modulating control (i.e., in lists that induce reactive control); in contrast, the relative weighting

scheme is shifted allowing more distal events to modulate control in high-probability conflict situations (i.e., in lists that induce proactive control).

Analysis 1 also indicated that the slope corresponding to the list-wide PC 50 condition was similar to the list-wide mostly congruent condition, reflecting a relatively short timescale of control unlike the list-wide mostly incongruent condition. Interestingly, this pattern mirrors a pattern Aben et al. (2017) found using a PC-50 volatile list condition. We reserve further discussion of this finding for the General Discussion.

Most importantly, the findings of Analysis 1 demonstrated that the timescale patterns observed by Aben et al. (2017) in list-wide mostly congruent and mostly incongruent conditions using a flanker task generalize to a different control-demanding task (picture-word Stroop) that utilized a design with diagnostic items that confirmed that differences between lists were attributable to differences in control. A caveat to note, however, is that the tasks used in Aben et al. (2017)
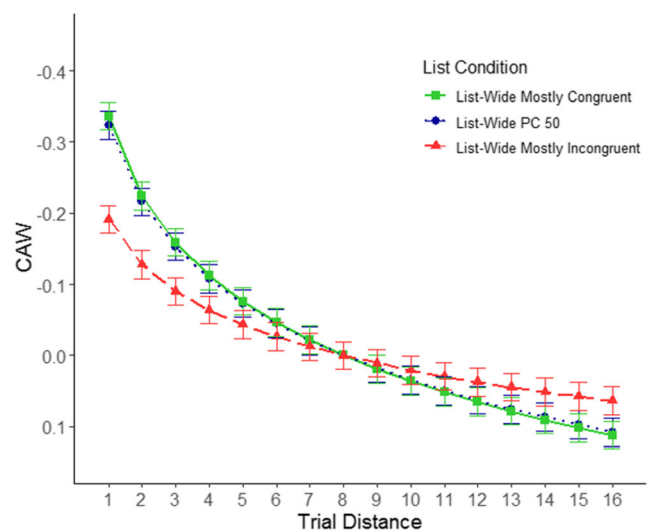
Table 2 Gonthier et al. (2016) List-Wide Proportion Congruence Full Model Coefficients

Variable	B (SE)	t
(Intercept)	-0.09 (0.02)	-4.74
Log trial distance	0.16 (0.02)	6.62
LWMC condition	-0.10 (0.03)	-3.67
LWMI condition	0.01 (0.03)	0.26
Log trial distance x LWMC condition	0.01 (0.03)	0.19
Log trial distance x LWMI condition	-0.06 (0.03)	-1.91

LWMC = list-wide mostly congruent; LWMI = list-wide mostly incongruent. The list-wide PC 50 condition was used as the reference condition (intercept).

Fig. 4 Estimates of the CAWs using inverse transformed RT (i.e., -10000/RT) from Model 4 (full model) in Analysis 1. Data are from a picture-word Stroop task (Gonthier et al., 2016). Each condition is plotted after subtracting its intercept. The original scale of trial distance is displayed on the x-axis, not log trial distance (which was included in the model for statistical testing). Error bars are +1/−1 standard deviation of the model predicted values. See Figure S1 in Supplementary Materials for estimates of CAWs as a function of log trial distance.

(flanker) and Analysis 1 (Gonthier et al., 2016) (picture-word Stroop) share a common feature in that they both entail stimuli for which the relevant and irrelevant dimensions are spatially separated. Thus, the timescale properties may be exclusive to tasks for which interference is at least partially controlled via spatial attention as opposed to more general attention mechanisms. An open question is whether the extended-CSE model would yield similar timescale patterns for an even more qualitatively different conflict task, such as a task for which the dimensions are fully spatially integrated (e.g., color word Stroop; Spieler, Balota, & Faust, 2000; see also Aschenbrenner & Balota, 2017). If the timescale of control is a task-general meta-control property, and the extended-CSE model is appropriately indexing timescales, then the model ought to result in similar timescale patterns across qualitatively different control tasks. If the timescale patterns are not replicable across tasks, then the only way to salvage the interpretation that the extended-CSE model captures timescales of control is by adopting a less parsimonious model of control that assumes different meta-control properties for different tasks. Thus, in Analysis 2, we investigated whether we would observe similar patterns when the model was applied to the color-word Stroop task.

Analysis 2: Examining the List-Wide Timescale of Control in a Color-Word Stroop Task from Gourley, Braver, and Bugg (2016)

For Analysis 2, we applied the extended-CSE model to an unpublished dataset (Gourley et al., 2016) from a color-word Stroop task in which list-wide proportion congruence was manipulated and participants named the ink color of a word while ignoring the word. Following the criteria specified in Footnote 1, we elected to use this dataset because it included a large number of participants ($N = 96$), the key factors were manipulated within-subjects, and diagnostic (PC 50) items were included in the task and the behavioral results from these items provided evidence that the list-wide proportion congruence manipulation induced differences in control.[6] That is, the same diagnostic items embedded in the mostly incongruent list yielded a smaller Stroop effect (93 ms) compared with when they were embedded in the mostly congruent list (107 ms). The color-word Stroop task is qualitatively different from both the flanker task and the picture-word Stroop task, because the relevant and irrelevant dimensions are spatially integrated (as opposed to fully separated in the flanker task and partially separated in the picture-word Stroop task; MacLeod, 1998; Spieler, et al. 2000). If the pattern of results is similar in a color-word Stroop task, this would

reinforce the conclusion that the timescale of control, as assessed by the extended-CSE model, is a task-general meta-control property of the control system.

Method

As was the case in the study from Analysis 1, in Gourley et al. (2016) each participant also was exposed to three conditions. They comprised a list-wide mostly congruent condition, a list-wide mostly incongruent condition, and a list-wide PC 60 condition. All three conditions included four diagnostic PC 50 items. What differed across conditions was the PC of the four other (inducer) items included in each list. In the list-wide mostly congruent condition, the other four items were PC 75. This led to an overall PC 67 list. In the list-wide mostly incongruent condition, the other four items were PC 25. This created an overall PC 33 list. The list-wide mostly congruent and incongruent conditions were comprised of 288 trials. For the third condition, the list-wide PC 60 condition,[7] the other four items were comprised of two PC 100 items and two PC 25 items. The latter two sets of items were used to manipulate item-specific probabilities of conflict (item-specific proportion congruence) within this list while more closely matching the global PC to the list-wide mostly congruent condition. The list-wide PC 60 condition contained 480 trials.

Analytical approach

There were 96 participants included in this dataset, and all were retained for the present analysis. As in Analysis 1, the analysis included all items (inducer and diagnostic). The following trials were excluded: the first trial of each condition (0.28%), error trials (3.67%), and trials following errors (3.67%). In addition, we box-trimmed the data such that trials <200 ms and >3,000 ms were excluded (0.82%). The RTs were transformed in the same fashion as in Analysis 1.

We used the same initial procedure as we did in Analysis 1 to determine how many trials back to include in the model. The results of this analysis are shown in Fig. 5. Based on these results, for the timescale of control analysis, we included up to eight trials back given that the t-values were above 1.96 through trial eight.

For the timescale of control analysis, RTs were modelled at two hierarchically related levels. The level one equation of the model was:

$$RT = \beta_0 + \beta_1 C_i + \beta_2 C_{(i-1)} + \beta_3 C_{(i-2)} \ldots + \beta_9 C_{(i-8)}$$
$$+ \beta_{10} C_i C_{(i-1)} + \beta_{11} C_i C_{(i-2)} \ldots + \beta_{17} C_i C_{(i-8)} \quad (3)$$

---

[6] Analysis of behavioral data are reported in more detail in the supplementary materials.

[7] The four diagnostic items in the PC 60 list condition yielded a Stroop effect of 110 ms.

Equation 3 was applied iteratively to each subject within each condition (list-wide mostly congruent condition, list-wide PC 60 condition, list-wide mostly incongruent condition). In doing so, a CAW for each of the previous eight trials was extracted for each subject within each condition. The subsequent procedure at the second level was identical to that of Analysis 1.

Results

The summary of the model comparisons is shown in Table 3. The test of Model 1 versus Model 0 indicated that CAWs significantly differed as a function of log trial distance, $X^2(1) = 127.48$, $p < 0.001$. The test of Model 2 versus Model 0 indicated that CAWs also significantly differed as a function of list condition, $X^2(2) = 12.44$, $p = 0.002$. The model with both log trial distance and list condition explained the data better than the model with just log trial distance, $X^2(2) = 13.18$, $p = 0.001$ (Model 3 vs. Model 1), and the model with just list condition, $X^2(1) = 128.22$, $p < 0.001$ (Model 3 vs. Model 2). Importantly, there was a significant difference when comparing the full model, which included the interaction of log trial distance and list condition (Model 4), and the main effects only model (Model 3), $X^2(2) = 19.74$, $p < 0.001$. As in Analysis 1, this warranted further analysis as to how the interaction terms (or slopes of the list conditions over trial distance) differed.

The regression coefficients of the full model are displayed in Table 4. Figure 6 displays the estimates of the full model corrected for the intercept of each condition. Similar to Analysis 1, we subsequently ran a no-intercept model to
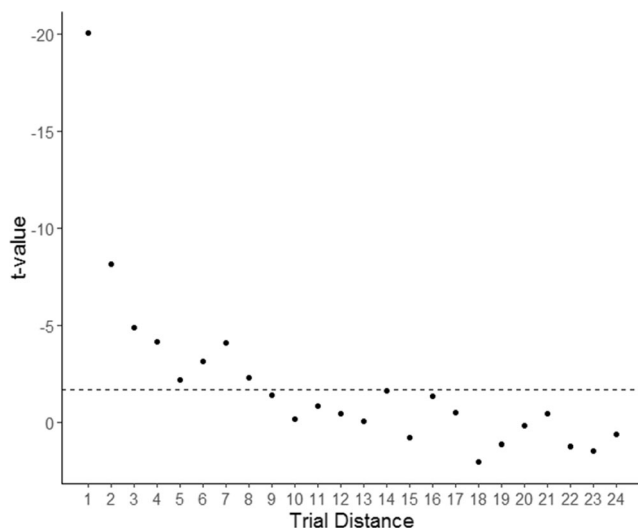

Fig 5 C_i C_{i-k} interaction t values from Eq. 3 plotted as a function of trial distance (ranges from 1-24). The dotted line represents the critical two-tailed t-value of 1.96 given infinite degrees of freedom. Data are from list-wide manipulations of a color-word Stroop task from Gourley et al. (2016).

compare timescales between conditions. The list-wide mostly congruent condition had a significantly shorter timescale relative to the list-wide mostly incongruent condition, $z = 2.66$, $p = 0.016$. We also observed a shorter timescale for the list-wide PC 60 condition relative to the list-wide mostly incongruent condition, $z = 4.73$, $p < 0.001$. Interestingly, the list-wide PC 60 condition also exhibited a shorter timescale relative to the list-wide mostly congruent condition, $z = 2.08$, $p = 0.038$.

Discussion

The results of Analysis 2 using data from a color-word Stroop task mostly align with the results from Aben et al. (2017) and Analysis 1. Most importantly, a shorter timescale of control was observed for the list-wide mostly congruent condition and a longer timescale of control was observed for the list-wide mostly incongruent condition. Interestingly, and differing from the PC 50 list in Analysis 1, the PC 60 list had a shorter timescale than the list-wide mostly congruent condition (see General Discussion for further consideration of the PC 60 timescale).

Given the general concordance of the results for list-wide mostly congruent and mostly incongruent conditions from Analyses 1 and 2 and Aben et al. (2017), the extended-CSE model appears to be robust in its ability to capture timescale patterns. In addition, the similarity of timescale patterns across tasks supports the possibility that timescales are a meta-control property of a task-general control mechanism.

Analysis 3: Examining the Item-Specific Timescale of Control in a Picture-Word Stroop Task from Bugg and Dey (2018).

Analyses 1 (picture-word Stroop) and 2 (color-word Stroop) demonstrated the generality of the timescale patterns that Aben et al. (2017) first observed in the context of a flanker task, patterns that Aben et al. attributed to the type of control (transient vs. sustained) engaged in mostly congruent compared with mostly incongruent lists, respectively (i.e., lists previously associated with reactive and proactive control; De Pisapia & Braver, De Pisapia & Braver, 2006). In short, with respect to the question of how far back the control system "looks" to learn about the probability of conflict within a list, the general answer was that it looks farther back in a mostly incongruent list than a mostly congruent list. In Analysis 3 we aimed to further test the generality of the timescale patterns by examining whether similar differences are observed for mostly congruent and mostly incongruent items.

This comparison, like the comparison of timescales of control for mostly congruent and mostly incongruent lists, similarly entails a contrast between conditions that differ in their probability of conflict and their corresponding behavioral effects.

Table 3 Gourley et al. (2016) List-Wide Proportion Congruence Model Comparisons

Model	df	AIC	log lik.	Test	X^2	p
0. Intercept	3	6104	-3049			
1. Log trial distance	4	5978	-2985	1 vs. 0	127.48	<0.001
2. Condition	5	6095	-3042	2 vs. 0	12.44	0.002
3. Log trial distance + Condition	6	5969	-2979	3 vs. 1	13.18	0.001
				3 vs. 2	128.22	<0.001
4. Log trial distance + Condition + Log trial distance x Condition	8	5953	-2969	4 vs. 3	19.74	<0.001

AIC = Akaike Information criterion; log lik. = log likelihood.

As noted earlier, the item-specific proportion congruence effect is the pattern whereby the Stroop effect is smaller for mostly incongruent items than mostly congruent items that are intermixed within a 50% congruent list. However, the comparison of timescales for mostly congruent and mostly incongruent items does not represent a comparison of conditions that differ in the type of control engaged. Rather, in both cases, reactive control is transiently recruited post-stimulus onset based on episodic associations with each item (Bugg, 2015; Bugg & Dey, 2018; Bugg & Hutchison, 2013; Bugg et al., 2011a; Chiu et al., 2017; cf. episodic retrieval account of Crump & Milliken, 2009). Presentation of a mostly congruent item reactively retrieves prior episodes that include an associated control setting that more fully processes the irrelevant dimension, whereas presentation of a mostly incongruent item reactively retrieves prior episodes that include an associated control setting that minimizes processing of the irrelevant dimension. Consequently, on the view that differences in timescales reflect differences in the type of control engaged (Aben et al., 2017), there should not be timescale differences between mostly congruent and mostly incongruent items. However, to date no study has examined whether the difference in performance between mostly congruent and mostly incongruent items may in part reflect differences in window size, that is, how far back the

control system looks to learn about the probability of conflict for each item. It may be that a smaller window size, corresponding to a shorter timescale, is characteristic of mostly congruent items whereas a larger window size, corresponding to a longer timescale is characteristic of mostly incongruent items. In other words, participants may weight recent experiences with mostly congruent items relatively more when adjusting control in response to the presentation of a mostly congruent item. Conversely, they may weight distal experiences with mostly incongruent items relatively more when adjusting control in response to the presentation of a mostly incongruent item. On this view, there should be timescale differences between mostly congruent and mostly incongruent items.

In Analysis 3, the extended-CSE model was applied to data from Bugg and Dey (2018). Akin to Gonthier et al. (2016), this study used an animal picture-word Stroop task to

Table 4 Gourley et al. (2016) List-Wide Proportion Congruence Full Model Coefficients

Variable	B (SE)	t
(Intercept)	-0.33 (0.03)	-9.89
Log trial distance	0.47 (0.05)	9.72
LWMC condition	0.02 (0.04)	.382
LWMI condition	0.13 (0.04)	2.99
Log trial distance x LWMC condition	-0.14 (0.07)	-2.04
Log trial distance x LWMI condition	-0.30 (0.07)	-4.45

LWMC = list-wide mostly congruent; LWMI = list-wide mostly incongruent. The list-wide PC 60 condition was used as the reference condition (intercept).
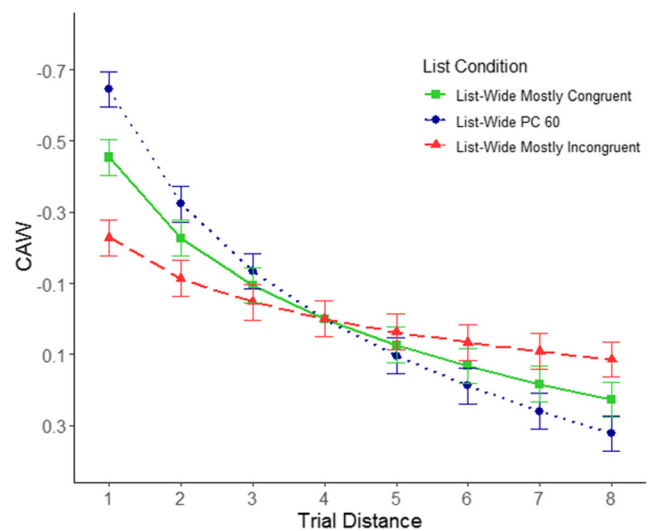

Fig. 6 Estimates of the CAWs using inverse transformed RT (i.e., −10,000/RT) from Model 4 (full model) in Analysis 2. Data are from a color-word Stroop task (Gourley et al., 2016). Each condition is plotted after subtracting its intercept. The original scale of trial distance is displayed on the x-axis. Error bars are +1/−1 standard deviation of the model predicted values. See Figure S2 in Supplementary Materials for estimates of CAWs as a function of log trial distance.

investigate cognitive control, namely item-level control. However, departing from Gonthier et al. (2016), all participants received a single list condition, which was PC 50. Importantly, this PC 50 list was comprised of mostly congruent and mostly incongruent items. As was the case in all other datasets discussed, data were available from a large sample[8], manipulations were within-subjects, and a confound-minimized design was employed (Braem et al., 2019).

A modification to the extended-CSE model was made to assess timescale differences between mostly congruent and mostly incongruent items within the PC 50 list. For item-level control, conflict learning reflects an accumulation of conflict experiences across mostly congruent and mostly incongruent *instances* of items as opposed to experiences across each successive trial. In other words, the random presentation of mostly congruent and mostly incongruent items within a list means that the previous trial's conflict status is not necessarily informative for investigating item-specific timescales of control (Fig. 7). Only the conflict statuses of previous instances of that item would be informative. Consequently, if there are differences in the timescale of control for items with different proportion congruences, then differences in the CAW slopes would be seen in previous *instances* of an item. Thus, the extended-CSE model was modified to use log *instance* distance (as opposed to log trial distance) as a predictor.[9]

Method

In Bugg and Dey (2018), there were four picture items (bird, cat, dog, and fish). These items were split into two overlapping sets (birds and cats; dogs and fish). One item set was mostly congruent and the other set was mostly incongruent (counterbalanced). The mostly congruent items were PC 67 and the mostly incongruent items were PC 33 such that the overall list was PC 50. Diagnostic items were also included,

and the pattern of results was consistent with confound-minimized item-level control.[10] Participants were presented with 432 trials. Additional details on the design and stimuli can be found in Bugg and Dey (2018).

Analytic approach Aggregating across the five experiments, there were 216 participants who were included in this analysis. The analysis included all items including diagnostic items. The following trials were excluded: the first trial of each condition (0.19%), error trials (4.79%), and trials following errors (4.79%). In addition, the data were box-trimmed such that trials <200 ms and >3,000 ms were excluded (0.39%). The RTs were transformed in the same fashion as in Analyses 1 and 2. We elected to include to up to six instances back in order to model timescales for mostly congruent and mostly incongruent items.[11] The resulting level one linear model for the timescale analysis was:[12]

$$RT = \beta_0 + \beta_1 C_i + \beta_2 C_{i-1} + \beta_3 C_{i-2} \ldots + \beta_7 C_{i-6}$$
$$+ \beta_8 C_i C_{i-1} + \beta_9 C_i C_{i-2} \ldots + \beta_{13} C_i C_{i-6} \quad (4)$$

C_{i-1} refers to the first prior instance of a given item (e.g., the last time a picture of a bird was presented). Equation 4 was applied iteratively to each subject for each item type (mostly congruent and mostly incongruent). In doing so, CAWs for each of the previous six instances of an item were extracted for each subject and for each item type (mostly congruent or mostly incongruent).

At the second level, the six CAWs estimated by Equation 5 were entered in as dependent variables in an HLM with log instance distance and item type as predictors. The predictors in the second level were entered in a stepwise fashion and each model was tested against its previous simpler nested model.

---

[8] We merged data from five experiments in Bugg and Dey (2018, Exp 1, Exp 2, Exp 3b, Exp 3c, and Exp 4b) resulting in a large sample of 216 participants. We included these experiments in the analysis because the Stroop phase of these experiments was identical and the behavioral results in RTs for all experiments yielded a reliable and typical item-specific proportion congruence effect such that the mostly congruent items had a larger Stroop effect than the mostly incongruent items, which transferred to diagnostic items.

[9] Note that we are not suggesting that log trial distance could not affect CAWs in the PC 50 list in Analysis 3. Indeed, Analysis 1 showed such an effect in a PC 50 list that, like the present PC 50 list, also had an item-specific proportion congruence manipulation embedded. As shown in Table S8 in Supplementary Materials, log trial distance also affected CAWs in the lists in Analysis 3. However, in contrast to Analysis 1, there are no other list types to compare this effect to in Analysis 3.

[10] In Bugg and Dey (2018) picture stimuli could either have been encoded as members of a category of pictures (e.g., bird, cat, dog, or fish) or individual exemplars (i.e., based on stimulus-specific perceptual features). PC 50 exemplars were presented as diagnostic items and results for these items revealed that items were encoded as members of categories. As such, in this analysis we treat the category as the item.

[11] Paralleling Aben et al. (2017) and Analyses 1 and 2, we used the significance level of previous instances as an inclusion criterion for the subsequent steps of the model. In Analysis 3, this yielded the inclusion of only three instances back in the subsequent steps of the model. We extended the number to six to more closely approximate Analyses 1 and 2. Subsequently, we confirmed that if we limited the analysis to three instances back as opposed to six, it did not change the results reported herein or our interpretations. Additionally, we confirmed that extending the number of prior instances (up to 12 to more closely approximate Analyses 1 and 2) also did not change the pattern of results or interpretations reported herein.

[12] As per the suggestion of a reviewer, we also assessed the effects of including the congruency state of previous trials and their interaction with the current trial in the model (i.e., the predictors used in Analyses 1 and 2). Thus, we modified Eq. 4 to include additional regressors that coded for congruency of previous trials and their interaction with the congruency state of the current trial. The resulting modification did not dramatically change our results nor our interpretations. The only change was that when we included previous trials and their interactions, the main effect of instance distance was marginally significant ($p = 0.099$), whereas previously (as reported in the main text) this main effect was significant ($p < 0.001$). Importantly, the interaction of instance distance and item type remained nonsignificant with ($p = 0.289$) or without ($p = 0.522$) these additional regressors. Details of this analysis can be found in Supplementary Materials.
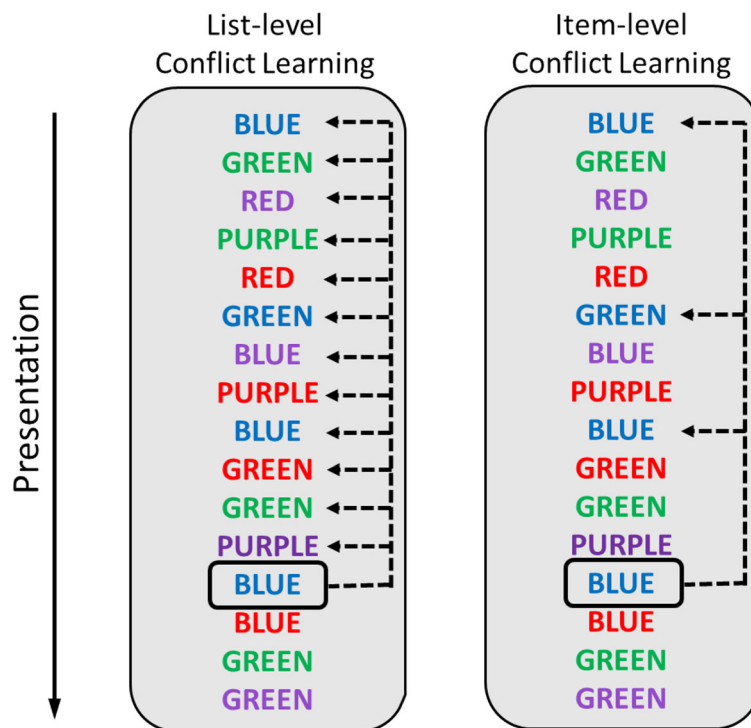
Fig. 7 Schematic representation of list-level conflict learning and item-level conflict learning. Items are presented sequentially from top to bottom. For list-level conflict learning, control is modulated by aggregating the conflict statuses of previous trials to arrive at a probability of conflict for the current circled trial. For item-level conflict learning, to modulate control for the current circled item (i.e., the color *blue*), the conflict statuses only of previous instances of that item must be aggregated.

Results

The summary of the model comparisons is shown in Table 5. The test of Model 1 versus Model 0 indicated that CAWs significantly differed as a function of log instance distance, $X^2(1) = 35.94$, $p < 0.001$. The test of Model 2 versus Model 0 indicated that CAWs did not differ as a function of item type, $X^2(2) = 0.25$, $p = 0.618$. Adding in item type with log instance distance did not explain the data better than the model with just log instance distance, $X^2(2) = 0.25$, $p = 0.615$ (Model 3 vs. Model 1). However, adding in log instance distance with item-type did explain the data better than the model with just item type, $X^2(1) = 35.94$, $p < 0.001$ (Model 3 vs. Model 2). Importantly, there was no significant difference when

comparing the full model, which included the interaction of log instance distance and list condition (Model 4), and the main effects only model (Model 3), $X^2(2) = 0.41$, $p = 0.522$.

The regression coefficients of the full model are displayed in Table 6. Figure 8 displays the estimates of the full model corrected for the intercept of each condition. The t-value for the interaction shows no significant difference between the slopes of the item types.

Discussion

Analysis 3 indicated no difference in the timescale of control for mostly congruent and mostly incongruent items within a PC 50 list. Thus, for items within a list that reactively trigger

Table 5 Bugg and Dey (2018) Item-specific Proportion Congruence Model Comparisons with Inverse Transformed RT

Model	df	AIC	log lik.	Test	X^2	p
0. Intercept	3	7320	-3657			
1. Log instance distance	4	7287	-3639	1 vs. 0	35.94	<0.001
2. Item type	4	7322	-3657	2 vs. 0	.25	0.618
3. Log instance distance + Item type	5	7288	-3639	3 vs. 1	.25	0.615
				3 vs. 2	35.94	<0.001
4. Log instance distance + Item type + Log instance distance x Item type	6	7290	-3639	4 vs. 3	.41	0.522

AIC = Akaike Information criterion; log lik. = log likelihood.

Table 6 Bugg and Dey (2018) Item-Specific Full Model Coefficients with Inverse Transformed RT

Variable	B (SE)	t
(Intercept)	-0.17 (0.03)	-5.50
Log instance distance	0.17 (0.04)	3.80
MI item	0.02 (0.04)	0.51
Log instance distance x MI item	0.04 (0.06)	0.64

MI = mostly incongruent. Mostly congruent items were used as the reference condition (intercept).

Figure 8 Estimates of the CAWs using inverse transformed RT (i.e., −10,000/RT) from Model 4 (full model) in Analysis 3. Data are from mostly congruent and mostly incongruent items within PC 50 lists from Bugg and Dey (2018). Each condition is plotted after subtracting its intercept. The original scale of instance distance is displayed on the x-axis. Error bars are +1/−1 standard deviation of the model predicted values. See Figure S3 in Supplementary Materials for estimates of CAWs as a function of log instance distance.

control, the timescale of control does not differ as a function of an item's probability of conflict. This is consistent with the notion that differences in the timescales of control for mostly congruent and mostly incongruent conditions are attributable to differences in the type of control (reactive vs. proactive) recruited in these conditions, as Aben et al. (2017) suggested. Furthermore, this finding indicates that participants are not differentially weighting past experiences with mostly congruent and mostly incongruent items (prior instances), for example by weighting recent or distal experiences with a given item type relatively more or less when adjusting control in response to the presentation of that item type. Rather, the finding suggests that the system looks equally far back (weights prior instances equally) when adjusting control for mostly congruent and mostly incongruent items.

General Discussion

The set of analyses presented in this paper aimed to accomplish two goals. One goal was to determine whether timescale patterns for list-wide proportion congruence manipulations observed in Aben et al. (2017) could be replicated in other conflict tasks using their extended-CSE model with the overarching aim of determining whether the model indexes a meta-control property of a task-general control mechanism. A second goal was to extend the model of Aben et al. to attempt to describe the timescale of control for items that varied in proportion congruence within a list (i.e., to document timescales for item-specific proportion congruence manipulations).

Collectively, the three analyses revealed several novel findings which converge on two major conclusions. One major conclusion is that timescale patterns of list-level control are replicable across qualitatively different conflict tasks, thereby supporting the view that the extended-CSE model indexes a task-general meta-control property. Supporting this conclusion are the results of Analyses 1 and 2 that evidenced differences in the slopes of the CAW trajectories for list-wide mostly congruent and list-wide mostly incongruent conditions in picture-word Stroop and color-word Stroop tasks, respectively. In both
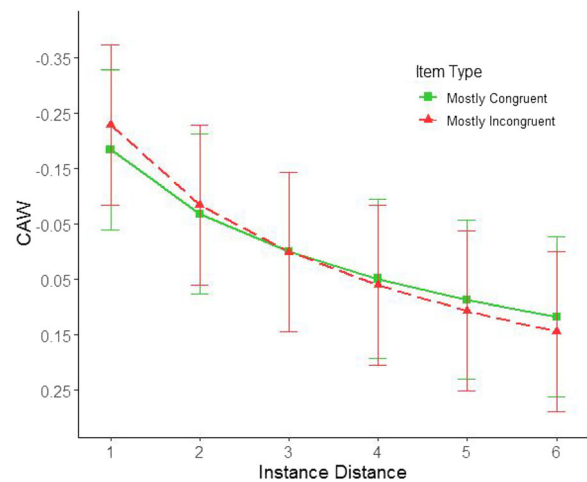
cases, the list-wide mostly congruent condition was characterized by steeper slopes, indicating a short timescale of control, and the list-wide mostly incongruent condition was characterized by less steep slopes, indicating a long timescale of control. While these patterns are consistent with Aben et al. (2017) who used a flanker task, two unique strengths of the present study bear mention. First, the present findings significantly extend those of Aben et al. because they were observed in tasks that implemented confound-minimized designs that warrant conclusions about control mechanisms as opposed to other mechanisms that can masquerade as control when such confounds are not controlled (e.g., contingency learning; Bugg et al., 2008; Blais & Bunge, 2010).[13] Second, our conclusions about timescales of list-level control are based on two qualitatively different conflict tasks, which allows us to infer that the timescales reflect a task-general characteristic of conflict tasks and not a task-specific one. This is further reinforced by the similarity of our findings to Aben et al. who used a different task than we used.

The second major conclusion of the present study informs the question of *why* these timescale patterns are observed. Collectively, our results suggest that the key factor underlying

---

[13] A reviewer appropriately pointed out that while the designs were confound-minimized due to the inclusion of diagnostic items, the timescale analyses used *both* inducer and diagnostic items to map out CAW trajectories. As such, while we can confidently assert that the list-wide proportion congruence effects observed for diagnostic items in these data sets indicate differences in control between lists, the conclusions we draw about the timescales of control are not confound minimized. That is, one could argue that participants did engage reactive and proactive control in mostly congruent and mostly incongruent lists, respectively, but the timescale patterns could be influenced by item-control processes or item-specific contingencies.

differences in timescales between different conditions (e.g., lists) is differences in the type of control engaged, as Aben et al. suggested. The results of Analyses 1 and 2 are consistent with this interpretation in that they indicated that the meta-control system placed more weight on relatively recent events in mostly congruent lists, which induce transient/reactive control, but weighted distal events to a greater degree in mostly incongruent lists, which induce sustained/proactive control. However, one could have alternatively suggested that it was differences in the probability of conflict per se that explained the differing timescale patterns given that in the case of the list-wide proportion congruence manipulation, type of control and probability of conflict are conflated. The results of Analyses 3 further support an interpretation based on differences in the type of control. In Analysis 3, we compared timescales between items that vastly differed in their probability of conflict (resulting in behavioral differences between items indicative of item-level control), yet no differences in timescales were observed. Mostly congruent items expressed the same timescale as mostly incongruent items. Because mostly congruent and mostly incongruent *items* both reactively trigger adjustments to attention, this finding is consistent with the view that the timescale of control is contingent on the type of control engaged (reactive vs. proactive; Aben et al., 2017).

Yet another finding that lends credence to this view is the timescale results for the list-wide PC 50 condition in Analysis 1. Recall that this list produced a timescale that was similar to the timescale produced by the list-wide mostly congruent condition even though the lists differed in their list-wide probability of conflict. We suspect that this reflects that reactive control was used both in the mostly congruent list, as previously discussed, and in the PC50 condition, which had an item-specific proportion congruence manipulation embedded within it such that PC 75 (mostly congruent), PC 50, and PC 25 (mostly incongruent) items were randomly intermixed.[14] Behavioral results indicated that participants modulated attention to these items differently (i.e., smaller Stroop effect for mostly incongruent items), indicative of reactive control. Interestingly, the shorter timescale observed in the list-wide PC 50 condition is reminiscent of the results observed in a volatile PC 50 list condition in Aben et al. (2017). In that condition the overall list was PC 50, but the list shifted between PC 80 and PC 20 every 20 trials (lists were 160 trials long). Aben et al. posited that the volatile condition likely induced greater reliance on reactive control because of its

unstable nature. Consistent with this view, they found shorter timescales for the volatile condition compared with a neutral PC 50 list condition in which PC did not change dramatically within the list (cf. Beherens et al., 2007; Jiang et al., 2014). Again, these patterns support the view that timescale differences reflect differences in the type of control and not differences in the probability of conflict between lists.

On this view, one result from Analysis 2 may at first appear surprising. In Analysis 2, we found that the list-wide PC 60 condition had a shorter timescale than the list-wide mostly congruent condition which was PC 67. Considering that the list-wide PC 60 condition (similar to the PC 50 condition in Analysis 1) biased use of reactive control (given the inclusion of items with different proportion congruencies within the list), as did the list wide mostly congruent condition, this may seem surprising. It also may seem surprising in light of the general pattern of results from Analysis 1 and Aben et al. (2017), indicating that a decrease in the list-level probability of conflict is associated with shorter timescales of control (which predicts a shorter timescale in the PC 67 list than the PC 60 list). One interpretation is that because the PC 60 condition used an item-specific manipulation, it encouraged even greater use of reactive control than the PC 67 list, which used a list-wide manipulation. In other words, there may be an additive effect of list-level conflict and item-specific manipulations for shifts to reactive control—the item-specific manipulation of the PC 60 list may have shortened the timescale of control enough to compensate for and "overcome" the 7% global conflict probability difference between the list-wide PC 67 list and the list-wide PC 60 list.

Taking stock, the present findings suggest that more engagement of reactive control processes leads to a relatively short timescale and more engagement of proactive control processes leads to a relatively long timescale. Regarding reactive control, this was true both in mostly congruent lists, where reactive control is triggered by interference demand and in lists in which an item-specific proportion congruence manipulation is embedded, where reactive control is triggered by episodic associations (Analyses 1 and 2). Mostly congruent items and mostly incongruent items within such lists, which both trigger reactive control adjustments, did not differ in their timescales (Analysis 3). This also supports the view that it is the type of control and not the probability of conflict per se that determines timescales, as do other patterns from the present study and Aben et al. (2017) as reviewed above. A novel prediction that falls out of this view is that individuals that heavily engage proactive control in a list-wide mostly congruent condition may express a long timescale of control, which would contradict the group-level patterns for list-wide mostly congruent conditions found in the present research but would be consistent with the view that the type of control and not probability of conflict per se is the key determinant of timescales. In the "Limitations and Future Directions" section

---

[14] Considering the definition of reactive control offered by Braver (2012), "reactive control reflects transient stimulus-driven goal reactivation…based on interference demands or episodic associations" (p. 106), in the case of the mostly congruent list reactive control was based on (unexpected) interference demands whereas in the case of the PC 50 list, where mostly congruent and mostly incongruent items were embedded, reactive control was based on episodic associations.

later below, we include discussion of models that include individual-level timescale-related parameters. Such models could plausibly test the above prediction.

Implications for Extant Models of Cognitive Control

The present findings mostly concur with the results from Aben et al. (2017) in demonstrating the ability to quantify timescales of control and by showing that meta-control timescale parameters systematically vary based on the type of control engaged. These findings provide a basis for updating extant conflict-modulated models of cognitive control (Blais et al., 2007; Botvinick et al., 2001; Verguts & Notebaert, 2008) that assume a fixed number of trials (or instances) back that a conflict monitoring unit utilizes to calculate the probability of conflict. In particular, the findings indicate that using a single learning rate parameter for a list-level conflict monitoring unit is not optimal, because low-probability conflict lists that induce reactive control produce short timescales, whereas high-probability conflict lists that induce proactive control produce long timescales. In other words, reactive control conditions seem to produce a large learning rate, such that the level of conflict is determined by relatively recent trials, and proactive conditions seem to produce a small learning rate such that the level of conflict is determined by both recent and distal trials. Notably, this is theoretically consistent with the dual mechanisms of control account (De Pisapia & Braver, 2006).

In addition, according to the item-level timescale results (Analysis 3), the timescale pattern for items does not follow the pattern for list-level control. That is, there are no differences in timescales based on items with differing probabilities of conflict. However, the findings of Analysis 1 and Analysis 3 together indicate that, within a PC 50 list, the congruency effect of the current trial is impacted both by previous trials (Analysis 1) *and* previous instances (Analysis 3 [see also Footnote 9]). Thus, it is possible that there are two conflict accumulators, one at the list-level (or at the global pathway level) and one at the item-level (Gonthier et al., 2016) and a hybrid model that includes both conflict accumulators may yield a more accurate representation of the control mechanism(s) at play (Blais et al., 2007; De Pisapia & Braver, 2006; Verguts & Notebaert, 2008).

Potential Neural Correlates of the Timescale of Control

Jiang et al. (2015) observed that variations in learning rates (timescales) at the participant level corresponded with variations in activity in the left anterior insula and inferior frontal gyrus. If these regions indeed represent timescales during control tasks,

then the present findings suggest that if we induced changes in control (e.g., proactive to reactive) we ought to observe corresponding online changes in these regions. For example, within a run, if we were to start with a list with a high probability of conflict and suddenly reduce the probability of conflict to a low level, we ought to observe corresponding changes in activity in the left anterior insula and inferior frontal gyrus.

In addition, it is widely assumed that being able to maintain a dynamic balance between different control types is an indication of a healthy control system (for review see Goschke & Bolte, 2014; Hommel, 2015). Assuming that a healthy meta-control system is sufficiently flexible and able to maintain a dynamic balance between short and long timescales, it is possible that indexing a measure of variability in the left anterior insula and inferior frontal gyrus may predict cognitive control ability. That is, if a person is able to dynamically shift between control types and consequently perform well in control tasks, this may be indexed by a measure of variability in these regions (see also, Burzynska et al., 2015; Garrett et al., 2011; Hu et al., 2014 for correlations between performance on attention tasks and fMRI BOLD variability).

Limitations and Future Directions

One limitation pertains to our conclusion that timescales of control may be a task-general meta-control property of the control system. Although the current findings in conjunction with those of Aben et al. (2017) support this conclusion, it is possible that the conclusion is specific to tasks for which conflict originates from stimulus processing (including Stroop and flanker) rather than response conflict (Simon; Egner, 2008; but see Hübner & Töbel, 2019 for an alternative view related to the similarity between flanker and Simon). A direct test of this possibility would be to examine timescales of control using manipulations that induce proactive and reactive control in a Simon task (see Hübner & Mishra, 2016; Toth et al., 1995). Relatedly, we must be clear that the current findings do not enable us to make any claims about whether the timescales reflect precisely the same underlying process(es) or region(s) of the brain (e.g., left anterior insula and inferior frontal gyrus as in Jiang et al., 2015) across tasks.

There are three additional limitations, which all relate to our analytic approach. First, Aben et al. (2017) inverse transformed RT to meet required assumptions for their modeling procedure, and we used the same transformation in our analyses because our goal was to determine whether their findings were replicable in other conflict tasks. However, the use of nonlinear transformations of RT data has been called into question, especially when investigating interactions (Balota et al., 2013; Lo & Andrews, 2015). Therefore, we also modeled the timescale of control with untransformed RT using generalized linear models with a Gamma link function

(see Supplementary Materials for more detail). The major resulting difference between the Aben et al. approach with transformed RT and the approach using untransformed RT was observed in Analysis 1. Specifically, when using untransformed RTs, the difference in the timescale of control between list conditions was not significant, although nominally the pattern of results was consistent. The results of Analyses 2 and 3 did not differ between the two approaches.

Second, the extended CSE model is more comprehensive than conventional CSE analyses because it includes more than a single previous trial when predicting behavior of a current trial. However, the number of parameters that are estimated by the level one equation is quite large. As a result, the CAWs for each person and each condition must be generated by individual ordinary least-squares regressions. At the second level of the HLM, these coefficients are not adjusted for the reliability of observations and are more susceptible to outliers (Snijders & Bosker, 2012). In comparison, in a more typical hierarchical linear modeling framework, the CAWs for each person in each condition would be adjusted to account for unreliable observations or outliers through shrinkage. A more typical hierarchical linear modeling procedure could not be applied to the extended-CSE model in the present set of analyses, because it would require many more trials per subject than what would be practically feasible to converge on a solution. In addition, each CAW parameter is assumed to have an *independent* effect on conflict modulation for the current trial. While the findings of Aben et al. (2017) and the current set of analyses provide evidence that distal trials do in fact independently contribute to control on current trials, it may be more appropriate to assume that conflict status is aggregated in some fashion across some window of trials and the joint contribution of previous trials impacts control. Alternative models have been specified that work around these issues by using learning rate parameters as a means of indexing the timescale of control (Chiu et al., 2017; Jiang et al., 2014; Jiang et al., 2015). The advantage of such models is that they greatly reduce the number of parameters to be estimated at the lowest level of a hierarchical linear model. Future efforts to model the timescale of control could draw inspiration from these models. Appropriately adjusted and potentially more accurate estimates may then be extracted from the same number of observations, and individual differences in timescales could be examined with a large enough sample size.

Third, there has been theoretical interest in whether CSEs are driven by effects of preceding incongruent trials or effects of preceding congruent trials (Berger et al., 2019; see Schlaghecken & Martini, 2012, for evidence that the absence of conflict is also a signal for control adjustments). However, the present study was not designed to address this question and we did not have sufficient trials to subdivide the interaction terms based on the preceding trial types.

Conclusions

In the present research, we applied an extended-CSE model developed by Aben et al. (2017) to two qualitatively different cognitive control tasks and demonstrated the model's ability to consistently capture the timescale for list-level control. The results replicated Aben and colleagues as we observed 1) a relatively short timescale for low conflict list-wide conditions (i.e., mostly congruent lists) and volatile list-wide conditions (PC 50 or PC 60 lists with item-specific proportion congruence manipulations), and 2) a relatively long timescale for high conflict list-wide conditions (i.e., mostly incongruent lists). These findings indicate that the extended-CSE model provides a fruitful means of quantifying timescales of control across conflict tasks and suggest that control processes that operate at the list-level share task general meta-control properties. Additionally, the present research found that a modified version of the extended-CSE model did not yield differences in the timescale of control at the item-level. Unlike mostly congruent and mostly incongruent lists, mostly congruent and mostly incongruent items were characterized by similar timescales of control despite their vastly differing probability of conflict and accompanying behavioral differences. The collective pattern of results in the present study, along with those of Aben et al., support the conclusion that the timescale of control depends on the type of control that is recruited (transient/reactive or sustained/proactive) and not on the probability of conflict per se that is associated with a given condition. These findings have implications for the refinement of mechanistic and theoretical models of cognitive control and meta-control.

References

Aben, B., Verguts, T., & Van den Bussche, E. (2017). Beyond Trial-by-Trial Adaptation: A Quantification of the Time Scale of Cognitive Control. *Journal of Experimental Psychology: Human Perception and Performance*.

Akaike, H. (1974). A new look at the statistical model identification. IEEE Transactions on Automatic Control 19(6), 716-723.

Aschenbrenner, A. J., & Balota, D. A. (2017). Dynamic adjustments of attentional control in healthy aging. *Psychology and Aging, 32*(1), 1-15.

Balota, D. A., Aschenbrenner, A. J., & Yap, M. J. (2013). Additive effects of word frequency and stimulus quality: The influence of trial

history and data transformations. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 39*(5), 1563-1571.

Beherens, T. E. J., Woolrich, M. W., Walton, M. E., & Rushworth, M. F. S. (2007). Learning the value of information in an uncertain world. *Nature Neuroscience, 10*, 1214–1221

Berger, A., Fischer, R. & Dreisbach, G. (2019). It's more than just conflict: The functional role of congruency in the sequential control adaptation. *Acta Psychologica, 197*, 64-72.

Blais, C., Robidoux, S., Risko, E.F., & Besner, D. (2007). Item-specific adaptation and the conflict monitoring hypothesis: A computational model. *Psychological Review, 114*, 1076–1086.

Blais, C., & Bunge, S. A. (2010). Behavioral and neural evidence for item-specific performance monitoring. *Journal of Cognitive Neuroscience, 22*, 2758–2767.

Blais, C., Harris, M. B., Guerrero, J. V., & Bunge, S. A. (2012). Rethinking the role of automaticity in cognitive control. *Quarterly Journal of Experimental Psychology, 65(2)*, 268–276.

Botvinick, M. M., Braver, T. S., Barch, D. M., Carter, C. S., & Cohen, J. D. (2001). Conflict monitoring and cognitive control. *Psychological Review, 108*, 624–652.

Braem, S., Bugg, J. M., Schmidt, J. R., Crump, M. J. C., Weissman, D. H., Notebaert, W., & Egner, T. (2019). Measuring Adaptive Control in Conflict Tasks. *Trends in Cognitive Sciences, 23*(9), 769–783.

Braver, T. S. (2012). The variable nature of cognitive control: A dual mechanisms framework. Trends in Cognitive Sciences, 16(2) , 106–113.

Braver, T.S., Gray, J.R., & Burgess, G.C. (2007). Explaining the many varieties of working memory variation: Dual mechanisms of cognitive control. In A.R.A. Conway, C. Jarrold, M.J. Kane, A. Miyake, & J.N. Towse (Eds.), *Variation in working memory* (pp. 76–106). Oxford University Press.

Bugg, J. M. (2012). Dissociating levels of cognitive control: The case of Stroop interference. *Current Directions in Psychological Science, 21*, 302 – 309.

Bugg, J. M. (2014). Conflict triggered top-down control: Default mode, last resort, or no such thing? *Journal of Experimental Psychology: Learning, Memory, and Cognition, 40*, 567 – 587.

Bugg, J. M. (2015). The relative attractiveness of distractors and targets affects the coming and going of item-specific control: Evidence from flanker tasks. *Attention, Perception, & Psychophysics, 77*, 373–389.

Bugg, J. M. (2017). Context, conflict, and control. In T. Egner (Ed.) The Wiley handbook of cognitive control. (pp. 79 – 96). Chichester, West Sussex: John Wiley & Sons.

Bugg, J. M., & Chanani, S. (2011). List-wide control is not entirely elusive: Evidence from picture-word Stroop. *Psychonomic Bulletin & Review, 18*, 930 – 936.

Bugg, J. M., & Crump, M. J. C. (2012). In support of a distinction between voluntary and stimulus-driven control: A review of the literature on proportion congruent effects. *Frontiers in Psychology, 3*, 367.

Bugg, J. M., & Hutchison, K. A. (2013). Converging evidence for control of color–word Stroop interference at the item level. *Journal of Experimental Psychology: Human Perception and Performance, 39(2)*, 433–449.

Bugg, J. M., & Dey, A. (2018). When stimulus-driven control settings compete: On the dominance of categories as cues for control. *Journal of Experimental Psychology: Human Perception and Performance*. Electronic publication ahead of print.

Bugg, J. M., Jacoby, L. L., & Chanani, S. (2011a). Why it is too early to lose control in accounts of item-specific proportion congruency effects. *Journal of Experimental Psychology: Human Perception and Performance, 37*, 844-859.

Bugg, J. M., Jacoby, L. L., & Toth, J. (2008). Multiple levels of control in the Stroop task. *Memory & Cognition, 36*, 1484–1494.

Bugg, J. M., McDaniel, M. A., Scullin, M. K., & Braver, T. S. (2011b). Revealing list-level control in the Stroop task by uncovering its benefits and a cost. *Journal of Experimental Psychology: Human Perception and Performance, 37,* 1595-1606.

Burzynska, A. Z., Wong, C. N., Voss, M. W., Cooke, G. E., McAuley, E., & Kramer, A. F. (2015). White matter integrity supports BOLD signal variability and cognitive performance in the aging human brain. *PLOS ONE, 10*(4), e0120315.

Chiu, Y. C., Jiang, J., & Egner, T. (2017). The caudate nucleus mediates learning of stimulus-control state associations. *Journal of Neuroscience, 37*, 1028 – 1038.

Cohen, J. D. (2017). Cognitive Control: Core Constructions and Current Considerations. In T. Egner (Ed.), *The Wiley handbook of cognitive control.* (pp. 17-41). Chichester, West Sussex, UK: John Wiley & Sons.

Cohen-Shikora, E. R., Suh, J., & Bugg, J. M. (2019). Assessing the temporal learning account of the list-wide proportion congruence effect. Journal of Experimental Psychology: Learning, Memory, and Cognition, 45(9), 1703–1723.

Colvett, J. S., Nobles, L. M., & Bugg, J. M. (2020). The unique effects of relatively recent conflict on cognitive control. *Journal of Experimental Psychology: Human Perception and Performance.* Advance online publication.

Crump, M. J. C., & Milliken, B. (2009). Short article: The flexibility of context-specific control: Evidence for context-driven generalization of item-specific control settings. *Quarterly Journal of Experimental Psychology, 62(8)*, 1523–1532.

De Pisapia, N., & Braver, T. S. (2006). A model of dual mechanisms through anterior cingulate and prefrontal cortex interactions. *Neurocomputing, 69,* 1322-1326.

Durston, S., Davidson, M. C., Thomas, K. M., Worden, M. S., Tottenham, N., Martinez, A., Watts, R., Ulug, A. M., & Casey, B. J. (2003). Parametric manipulation of conflict and response competition using rapid mixed-trial event-related fMRI. NeuroImage, 20(4), 2135–2141.

Duthoo, W., Abrahamse, E. L., Braem, S., Boehler, C. N., Notebaert, W. (2014). The Congruency Sequence Effect 3.0: A Critical Test of Conflict Adaptation. *PLoS ONE, 9*, e110462.

Egner, T. (2007). Congruency sequence effects and cognitive control. *Cognitive, Affective, & Behavioral Neuroscience, 7*, 380–390.

Egner, T. (2008). Multiple conflict-driven control mechanisms in the human brain. *Trends in Cognitive Sciences, 12*, 374–380.

Garrett, D. D., Kovacevic, N., McIntosh, A. R., & Grady, C. L. (2011). The importance of being variable. *Journal of Neuroscience, 31*(12), 4496–4503.

Gonthier, C., Braver, T. S., & Bugg, J. M. (2016). Dissociating proactive and reactive control in the Stroop task. *Memory & Cognition, 44*, 778 – 788.

Goschke, T., & Bolte, A. (2014). Emotional modulation of control dilemmas: The role of positive affect, reward, and dopamine in cognitive stability and flexibility. *Neuropsychologia, 62*, 403–423.

Gourley, E. M., Braver, T. S., & Bugg, J. M. (2016). Dissociating proactive and reactive control: A replication and extension using color-word Stroop. Poster presented at the 57th annual meeting of the Psychonomics Society, Boston, MA.

Gratton, G., Coles, M. G. H., & Donchin, E. (1992). Optimizing the use of information: Strategic control of activation of responses. *Journal of Experimental Psychology: General, 121*(4), 480-506.

Hommel, B. (2015). Between persistence and flexibility: The Yin and Yang of action control. In A. J. Elliot (Ed.), *Advances in Motivation Science* (Vol. 2, pp. 33–67).

Horga, G., Maia, T. V., Wang, P., Wang, Z., Marsh, R., & Peterson, B. S. (2011). Adaptation to conflict via context-driven anticipatory signals in the dorsomedial prefrontal cortex. Journal of Neuroscience, 31(45), 16208–16216.

Hothorn, T., Bretz, F., & Westfall, P. (2008). Simultaneous inference in general parametric models. *Biometrical Journal, 50,* 346-363.

Hu, S., Chao, H. H.-A., Zhang, S., Ide, J. S., & Li, C.-S. R. (2014). Changes in cerebral morphometry and amplitude of low-frequency fluctuations of BOLD signals during healthy aging: Correlation with inhibitory control. *Brain Structure and Function, 219*(3), 983–994.

Hübner, R., & Mishra, S. (2016). Location-specific attentional control is also possible in the Simon task. *Psychonomic Bulletin & Review, 23,* 1867–1872.

Hübner, R., & Töbel, L. (2019). Conflict resolution in the Eriksen flanker task: Similarities and differences to the Simon task. *PLOS ONE, 14,* e0214203.

Hutchison, K. A. (2011). The interactive effects of list-based control, item-based control, and working memory capacity on Stroop performance. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 37,* 851–860.

Jacoby, L.L., Lindsay, D.S., & Hessels, S. (2003). Item-specific control of automatic processes: Stroop process dissociations. *Psychonomic Bulletin & Review, 10*(3), 638–644.

Jiang, J., Heller, K., & Egner, T. (2014). Bayesian modeling of flexible cognitive control. *Neuroscience and Biobehavioral Reviews, 46,* 30–43.

Jiang, J., Beck, J., Heller, K., & Egner, T. (2015). An insula-frontostriatal network mediates flexible cognitive control by adaptively predicting changing control demands. *Nature Communications, 6,* 8165–8176.

Jiménez L., & Méndez A. (2013). It is not what you expect: dissociating conflict adaptation from expectancies in a Stroop task. *Journal of Experimental Psychology: Human Perception and Performance, 39,* 271–284.

Jiménez, L., & Méndez, A. (2014). Even with time, conflict adaptation is not made of expectancies. *Frontiers in Psychology, 5.*

Kane, M.J., & Engle, R.W. (2003). Working-memory capacity and the control of attention: The contributions of goal maintenance, response competition, and task set to Stroop interference. *Journal of Experimental Psychology: General, 132,* 47-70.

Kinoshita, S., Mozer, M. C., & Forster, K. I. (2011). Dynamic adaptation to history of trial difficulty explains the effect of congruency proportion on masked priming. *Journal of Experimental Psychology: General, 140,* 622-636.

Lindsay, D. S., & Jacoby, L. L. (1994). Stroop process dissociations: The relationship between facilitation and interference. *Journal of Experimental Psychology: Human Perception and Performance, 20,* 219–234.

Lo, S., & Andrews, S. (2015). To transform or not to transform: Using generalized linear mixed models to analyse reaction time data. Frontiers in Psychology, 6.

Logan, G. D. & Zbrodoff, N. J. (1979). When it helps to be misled: Facilitative effects of increasing the frequency of conflicting stimuli in a Stroop-like task. *Memory & Cognition, 7,* 166–174.

Logan, G. D., Zbrodoff, N. J., Williamson, J. (1984). Strategies in the color-word Stroop task. *Bulletin of the Psychonomic Society, 22,* 135–138.

Lowe, D., & Mitterer, J. O. (1982). Selective and divided attention in a Stroop task. *Canadian Journal of Psychology, 36,* 684–700.

Schmidt, J. R., & Besner, D. (2008). The Stroop effect: Why proportion congruence has nothing to do with congruency and everything to do with contingency. *Journal of Experimental Psychology: Learning, Memory, & Cognition, 34,* 514–523.

Schlaghecken, F., & Martini, P. (2012). Context, not conflict, drives cognitive control. *Journal of Experimental Psychology: Human Perception and Performance, 38*(2), 272.

Snijders, T. A., & Bosker, R. J. (2012). *Multilevel analysis: An introduction to basic and advanced multilevel modeling*. Los Angeles: SAGE.

Spieler, D. H., Balota, D. A., & Faust, M. E. (2000). Levels of selective attention revealed through analyses of response time distributions. Journal of Experimental Psychology: Human Perception and Performance, 26(2), 506–526.

Toth, J. P., Levine, B., Stuss, D. T., Oh, A., Winocur, G., & Meiran, N. (1995). Dissociation of processes underlying spatial S–R compatibility: Evidence for the independent influence of what and where. *Consciousness and Cognition, 4,* 483–501.

Verguts, T., & Notebaert, W. (2008). Hebbian learning of cognitive control: Dealing with specific and nonspecific adaptation. *Psychological Review, 115,* 518-525.

Open Practices Statement

The data or materials for the experiments reported here are available upon request to corresponding author. None of the analyses were preregistered.