Spring 5-2019

Learning from Past Conflict: Investigating the Time Scale of Conflict Learning for Cognitive Control Processes

Abhishek Dey
Washington University in St. Louis

Follow this and additional works at: https://openscholarship.wustl.edu/art_sci_etds

Part of the Cognitive Psychology Commons

WASHINGTON UNIVERSITY IN ST. LOUIS

Psychological and Brain Sciences Department

Learning from Past Conflict: Investigating the Time Scale of Conflict Learning for Cognitive
Control Processes
by
Abhishek Dey

A thesis presented to
The Graduate School
of Washington University in
partial fulfillment of the
requirements for the degree
of Master of Arts

May 2019
St. Louis, Missouri

Table of Contents

List of Figures

List of Tables

Acknowledgments

I acknowledge my mentor Dr. Julie Bugg for her insight and assistance in completing this project and for guiding my development as a scientist. I thank Dr. Todd Braver and Dr. David Balota for serving as members of my thesis committee. I thank members of the Cognitive Control and Aging Laboratory for their support and suggestions. I thank the Graduate School and the Psychological and Brain Sciences Department for their financial support.

Abhishek Dey

Washington University in St. Louis

May 2019

Special thanks to my parents.

Abstract of the Thesis

Learning from Past Conflict: Investigating the Time Scale of Conflict Learning for Cognitive

Control Processes

by

Abhishek Dey

Master of Arts in Cognitive Psychology

Washington University in St. Louis, 2019

Professor Julie Bugg, Chair

Conflict-modulated cognitive control accounts posit that control processes adjust attention based on the probability of conflict associated with a given context (e.g., list of items, a particular item within a list, etc.). However, within these accounts, it is not yet fully understood *how* the control system learns about the probability of conflict. A specific question I address in the present research is how far back does the control system look to learn about the probability of conflict? In other words, what is the time scale of conflict learning for the control system? I use a statistical model recently developed by Aben et al. (2017) that captured the time scale of conflict learning for list-level control processes in a flanker task. The set of analyses I present shows that this model reliably captures the time scale of conflict learning for task-general, list-level control processes (Analysis 1 and Analysis 2). In addition, I also demonstrate that there are no differences in the time scale of conflict learning for differentially conflicting items within a list which are thought to engage item-level control (Analysis 3 and Analysis 4). I discuss potential reasons for the time scale patterns and the implications they may have for extant theories of cognitive control.

Introduction

      Cognitive control, or attentional control, can be thought of as the ability to pursue goal-directed behavior in the face of more habitual or immediately compelling alternative behaviors (Cohen 2017). Cognitive control is often driven by changes in the probability of conflict in a given context (e.g., Bugg, 2014; Bugg, Jacoby, & Chanani, 2011; Crump, Gong, & Milliken, 2006; Jacoby, Lindsay, & Hessels, 2003; Logan & Zbrodoff, 1979). Conflict occurs when the relevant (to-be-attended) and irrelevant (to-be-ignored) stimulus dimensions trigger competing cognitive representations or response options. The term *context-driven control* implies that the control system learns about the probability of conflict within a given context and transforms that information into a signal that modulates attention appropriately (i.e., to match the probability). For example, imagine a student is attending two lecture classes that are held in different classrooms in the same building. In one class noise from construction intermittently interrupts class. The other class is not interrupted by noise. At the beginning of the semester, the student's ability to attend to the lecture may be impaired in the class with construction noise relative to the class with no noise. However, as a function of the control system learning over the course of the semester, the control system may continuously heighten attention towards the lecture and filter out distraction (e.g., noise) when in the classroom in which there is a high probability of noise. In comparison, in the class in which there are rarely any interruptions, the control system may not heighten attention in this fashion. Consequently, the student may find it easier to pay attention to the lecture in the classroom with the construction noise. This context-driven allocation of attention on a classroom-by-classroom basis demonstrates how the control system learns the statistical regularities of conflict based on accumulating experiences within a context (e.g., the

1

probability of conflict), and transforms the learned information to produce a signal that heightens attention settings for one class and not the other.

Several mechanistic models of cognitive control have been developed to explain how the control system modulates attention settings. In this paper, I focus on the many models that assume that the signal for control adjustment is conflict. These include (but are not limited to) the conflict-monitoring hypothesis (Botvinick, Braver, Barch, Carter, Cohen, 2001), the item-specific adaptation and conflict-monitoring hypothesis (Blais, Robidoux, Risko, & Besner, 2007), and the dual mechanisms of control account (De Pisapia & Braver, 2006). These models all agree that past experiences of conflict are central to how the system modulates attention settings. In addition, a strong interpretation of these models assumes that there is a task-general control system as opposed to task-specific control systems. As an everyday example, these models would imply that there is a single control system responsible for modulating attention to lectures in class and to the road when driving; in the laboratory, the implication is that there is a single control system for modulating attention to distractors in a flanker task and in a Stroop task. One outstanding question that remains within these accounts is how the control system learns about statistical regularities of conflict. Specifically, what is the time scale under which a conflict learning mechanism operates? Here, a conflict learning mechanism can be thought of as a conflict accumulator, that is, a mechanism that stores and represents prior instances of conflict. Time scale refers to how many prior trials are tracked by the accumulator. In reference to the earlier example, does the student's conflict accumulator aggregate conflict information across the whole semester (long time scale), or does the accumulator aggregate conflict information across just a few previous days (short time scale).

In an attempt to tackle the above outstanding question, Aben, Verguts and Van den Bussche (2017) recently developed a statistical model to quantify the time scale of conflict learning for contexts with differing probabilities of conflict. But, before I describe the model and results from Aben et al., I will provide some background that will serve as a foundation for understanding their results and for understanding the motivation behind the present research. First, I will begin by detailing a way that we can manipulate the global probability of conflict (i.e., proportion congruence) and I will describe a behavioral pattern that result as a function of this manipulation (i.e., the proportion congruence effect). This behavioral pattern serve as an initial motivator for examining potential differences in the time scale of conflict learning between contexts. Second, I will describe a way that we can manipulate probability of conflict for a given item and the behavioral pattern that result from such a manipulation. This will serve as motivation for examining time scale differences as a function of probability of conflict for items within a context.

Proportion Congruence Effects

In attention demanding tasks like Stroop, flanker, and Simon tasks, the congruency effect is the observation that participants are faster to respond to, and are sometimes more accurate on, congruent trials as compared to incongruent trials. For example, in color-word Stroop where participants name the ink color of color words, this is reflected in faster reaction times (RT) and higher accuracy for congruent stimuli (e.g., the word matches the color blue) compared to incongruent stimuli (e.g., the word **GREEN** conflicts with the color blue). Critically for present purposes, the congruency effect systematically varies as a function of global probability of conflict within a context. When the probability of conflict is low, as in a mostly congruent list,

the congruency effect is larger than when the probability of conflict is high, as in a mostly incongruent list (Logan & Zbrodoff, 1979). This robust behavioral pattern is termed the proportion congruence effect (Lindsay & Jacoby, 1994; Logan & Zbrodoff, 1979; Logan, Zbrodoff, & Williamson, 1984; Lowe & Mitterer, 1982). The prevailing interpretation of this pattern is that the control system learns about the statistical regularities of conflict within a list and biases attention settings based on those regularities (Lindsay & Jacoby, 1994; Logan & Zbrodoff, 1979; Lowe & Mitterer, 1982; Tzelgov, Henik, & Berger, 1992). A mostly congruent list is associated with a setting that processes the word to a greater extent than the setting associated with the mostly incongruent list (Lindsay & Jacoby, 1994; cf. Melara & Algom, 2003), in which attention to color may also be amplified (Egner & Hirsch, 2005). A natural question that follows from this interpretation, and a primary one this present research seeks to address, is how far back does the control system look to learn about the probability of conflict within a given list context? In other words, what is the time scale of conflict learning? As an exaggerated example, when reaching the 100th item in a Stroop list, has the conflict learning mechanism accumulated information from all previous 99 items? Or, is there a restricted range for accumulation such that only a few previous trials are aggregated to calculate probability of conflict, and thereby modulate attention? Moreover, does the time scale of conflict learning systematically vary as a function of global probability of conflict? That is, is there a difference in time scales between mostly congruent (low probability of conflict) and mostly incongruent (high probability conflict) contexts?

Conflict Learning at Multiple Levels of Control

A point to consider for developing models of the conflict learning mechanism is that conflict learning can occur at multiple levels. To make a coarse distinction, control can be applied globally across the whole context, or locally for items within a context. In the previous section, I described proportion congruence effects, but of a particular kind – *list-wide* proportion congruence effects. To reiterate, one account of this effect is that the control system learns about the global probability of conflict associated with an entire list. Because the probability of conflict of the list guides attention, control is thought to operate at the list-level (Bugg, 2014; Bugg & Chanani, 2011; Hutchison, 2011; Gonthier, Braver, & Bugg, 2016). However, with traditional list-wide proportion congruence manipulations, items within a list have item-specific proportion congruences that match the list-wide proportion congruence. That is, a color-word Stroop list with a proportion congruence of .75 (PC 75) is made up of items that are also PC 75. In other words, the color **green** is congruent on 75% of trials, as is the color **blue**, etc. Thus, an alternative account of list-wide proportion congruence effects is that the system learns about the probability of conflict for *items* within a list and guides attention using item-level control (Blais & Bunge, 2010; Blais et al., 2007). A way to adjudicate between these two accounts makes use of additional items termed *transfer items*. Transfer items are items that have an overall proportion congruence of .50 (PC 50). These items are embedded within mostly congruent and mostly incongruent lists. List-level control is demonstrated when PC 50 transfer items within a mostly congruent list have larger Stroop effects than the same PC 50 transfer items within mostly incongruent lists (see Bugg, 2014; Gonthier et al., 2016; Bugg & Chanani, 2011; see also Hutchison, 2011, for the same pattern with PC matched transfer items that were not 50% congruent).

While it is now clear that list-level control can be observed independent of item-level control, behavioral evidence also clearly demonstrates that item-level control can be observed independent of list-level control in certain contexts. Jacoby, Lindsay, and Hessels (2003) found that mostly congruent items within a list produced larger Stroop effects than mostly incongruent items within the same list, a pattern known as an *item-specific* proportion congruence effect. Importantly, these items were intermixed such that the list itself was PC 50. This effect has since been observed after controlling for contingency confounds in Stroop and flanker tasks (e.g., Bugg, 2015; Bugg & Dey, 2018; Bugg & Hutchison, 2013; Bugg, Jacoby, & Chanani, 2011; Chiu, Jiang, & Egner, 2017; for role of contingency learning, see Schmidt & Besner, 2008), demonstrating modulations of attention on an item-by-item basis. Because attention can be modulated for items within a list, the same questions regarding the time scale of conflict learning for list-level control can also be asked for item-level control. Namely, how long is the time scale of conflict learning for item-level control, and does the time scale vary systematically for mostly congruent and mostly incongruent items?

A framework which makes explicit the differences between item-level control and list-level control is the dual mechanisms of control account (Braver, Gray, & Burgess, 2007; see also Bugg, 2012; Bugg, 2017). The dual mechanisms of control account states that there are two separable control mechanisms – proactive control and reactive control. List-level control is thought to be proactive and is a tonic/sustained mechanism that does not fluctuate substantially trial-by-trial. As it relates to the previous discussion of transfer (PC 50) items; the reason differing congruency effects are observed for transfer items between mostly congruent and mostly incongruent lists is because a sustained attentional state based on the global probability of conflict for a list is utilized. In contrast, item-level control is thought to be a phasic/transient

6

mechanism such that a transient attentional state based on the probability of conflict for a given item is utilized. Because mostly congruent and mostly incongruent items are intermixed within a list, there is no way that the system can prepare in advance for an item. Thus, any control adjustments that are observed must arise after the item is presented.

Based on the dual-mechanisms of control account, proactive sustained control would be predicted to have a long time scale of conflict learning and reactive transient control would be predicted to have a short time scale of conflict learning. Following these predictions, in the present research I ought to observe long time scales for list-level control and short time scales for item-level control if they reflect proactive and reactive control mechanisms, respectively. However, an important finding as it relates to the dual mechanism of control account is that although both list-wide mostly congruent and list-wide mostly incongruent manipulations parametrically change conflict levels across the entire list, list-wide mostly congruent manipulations result in behavioral and neuropsychological markers that are more consistent with reactive control mechanisms and list-wide mostly incongruent manipulations result in markers that are more consistent with proactive control mechanisms (De Pisapia & Braver, 2006). Given these findings, the predictions would be that list-wide mostly incongruent manipulations lead to long time scales of conflict learning, whereas both list-wide mostly congruent and item-level control lead to short time scales of conflict learning.

I have thus far provided a brief summary of the relevant background literature that motivated my inquiry into the time scale of conflict learning. I have also made explicit predictions regarding the relative time scales of conflict learning for different types of control. In the next section I will describe a novel method used to quantify the time scale of conflict learning developed by Aben et al. (2017). The statistical model they used is central to the

present research as I used their exact modeling procedure in Analysis 1 and 2 and but modified the model for Analysis 3 and 4.

Quantifying the Time Scale of Conflict Learning

To fully appreciate the model developed by Aben and colleagues (2017), I will briefly describe another behavioral indicator of cognitive control that the authors leveraged – the congruency sequence effect (CSE). There has been much literature demonstrating that the congruency effect (i.e., the difference in reaction time between congruent and incongruent trials) is subject to cross-trial sequence effects such that conflict on the previous trial ($C_{i-1;}$) influences the degree to which conflict on the current trial (C_i) affects RT and accuracy (Gratton, Coles, & Donchin, 1992; for reviews, see Duthoo, Abrahamse, Braem, Boehler, & Notebaert, 2014; Egner, 2007). The impact of C_{i-1} on C_i (i.e, the $C_i C_{i-1}$ interaction) is termed the congruency sequence effect (CSE). Figure 1 depicts the typical pattern for the CSE whereby the congruency effect is reduced following a previous trial that is incongruent compared to one that is congruent.

One prominent interpretation of the CSE is that it reflects local adjustments of attention by the control system based on the detection of conflict (Botvinick, Braver, Barch, Carter, & Cohen, 2001). That is, based on the conflict status of the previous trial, the system adjusts attention to the next trial. When encountering conflict on trial i-1, the control system relatively increases attention to the relevant stimulus dimension reducing the congruency effect for the current trial. Of primary interest for present purposes, Aben et al. (2017) recently examined an extension of the CSE in which they looked not just at 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 of the current trial. 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. Here, k is equal to the maximum trial distance that had a significant impact on the current trial. The primary parameters of interest were the current trial x previous trial interactions (e.g., C_iC_{i-1}, C_iC_{i-2}, ….. C_iC_{i-k}). In this model, the value of an interaction term for a given trial distance is deemed to be that trial distance's conflict-adaptation-weight (CAW). That is, the CAW for a given trial distance quantifies that trial distance's *independent* impact on the current trial's congruency effect. In addition, the relative change in CAWs across trial distances (slopes in Figure 2) is interpreted as an index of the time scale of conflict learning. A long time scale is indexed by CAWs that do not change much over trial distance, whereas a short time scale is indexed by CAWs that change more dramatically over trial distance. A long time scale is reflective of a reduced learning rate of conflict. Essentially, it translates to a relative decrease in the weighting of recent trials and a relative increase in the weighting of distal trials when compared to a short time scale. Figure 2 (from Aben et al., 2017) illustrates the difference between long and short time scales as captured by their model. Henceforth I will refer to their model as the extended-CSE model.

Aben et al. (2017) used the extended-CSE model to quantify differences in the time scale of conflict learning between four different list-wide manipulations in a flanker task. The four list-wide manipulations were as follows – 1) list-wide mostly congruent, 2) list-wide mostly incongruent, 3) neutral, and 4) volatile. The mostly congruent list was PC 80, the mostly incongruent list was PC 20, the neutral list was PC 50, and the volatile list was also PC 50. The key difference between the neutral list and volatile list was that the PC of the volatile list shifted between PC 80 and PC 20 every 20 trials. Each list was 160 trials long. Based on the dual mechanisms of control account, and consistent with the predictions made in the present research, the authors predicted that the list-wide mostly incongruent manipulation would result in a long

time scale of conflict learning. In contrast, the list-wide mostly congruent manipulation was predicted to result in a short time scale of conflict learning. In addition, they predicted that the volatile list would result in a short time scale of conflict learning as there has been evidence showing that the conflict learning rate increases as probability of conflict changes rapidly across a list (Beherens, Woolrich, Walton, & Rushworth, 2007; Jiang, Heller, & Egner, 2014).

The Extended-CSE Modeling Procedure

In this section I will describe the extended-CSE model in more detail. To reiterate, in the extended-CSE model the interaction terms from the statistical model (CAWs) are interpreted as indexes of how much the first previous through k^{th} previous trial impacted the current trial's congruency effect. If there was a significant effect, then that trial was included as part of the window size for conflict accumulation. In order to determine a suitable window size for their time scale analysis, Aben and colleagues (2017) first chose a starting value of 14 for k. The authors then ran a linear mixed effects model (nested within subjects), across the entire sample and collapsing across lists, with RT as the dependent measure. The predictors were the congruency state of the current trial, C_i, (0 – congruent, 1 – incongruent), the congruency state of the first previous trial up to the 14th previous trial, C_{i-1}, C_{i-2}, ….. C_{i-14}, and the interactions between current trial congruency and preceding trial congruency, C_iC_{i-1}, C_iC_{i-2}, ….. C_iC_{i-14}. By running this model, they were able to determine how many trials back they should include to quantify the time scale of control for each condition. They found that the previous 12 trials had significant or close to significant interaction terms.

Using k = 12 for their time scale analysis, Aben and colleagues (2017) analyzed RTs at two levels. The first level used linear models for each subject in each condition to extract the interaction coefficients from the following equation:

$$RT = \beta_0 + \beta_1 C_i + \beta_2 C_{i-1} + \beta_3 C_{i-2} \ldots + \beta_{13} C_{i-12} + \beta_{14} C_i C_{i-1} + \beta_{15} C_i C_{i-2} \ldots + \beta_{25} C_i C_{i-12}$$

At the second level, they extracted the coefficients of the interaction terms and used those coefficients as dependent variables with log trial distance and condition as predictors. Consistent with their predictions, the authors found that the list-wide mostly congruent and volatile list conditions produced a shorter time scale of conflict learning relative to list-wide mostly incongruent and neutral list conditions. That is, list-wide mostly congruent and volatile conditions induced more rapid changes with more recent incongruent trials dominating CAWs, whereas list-wide mostly incongruent and neutral conditions showed slower and more steady changes in CAWs across trial distance. An interpretation of these results is that the conflict learning system in list-wide mostly congruent and volatile conditions utilizes fewer trials to make control adjustments relative to list-wide mostly incongruent and neutral conditions. Focusing just on the list-wide mostly congruent and list-wide mostly incongruent conditions, this is consistent with the predictions that fall out of the dual mechanisms of control account which states that reactive control (list-wide mostly congruent) ought to produce a short time scale and proactive control (list-wide mostly incongruent) ought to produce a long time scale.

Present Study

Motivated by the results from Aben et al. (2017), in the present research, I attempted to replicate the pattern they observed and extend their findings in two important ways. To replicate, I used a data set from a picture-word Stroop task from Gonthier, Braver, and Bugg

11

(2016). I chose to use data from the Gonthier et al. study because the design revealed list-level control. Specifically, the data showed that PC 50 transfer items within a mostly congruent list had larger Stroop effects than PC 50 transfer items within a mostly incongruent list. I applied the extended CSE-model to this data and investigated whether the time scale of conflict learning varies systematically based on global probability of conflict between lists (Analysis 1). To follow up, the two extensions I made were as follows. First, I investigated how generalizable the list-level time scale pattern is across different control demanding tasks. I applied the extended-CSE model to a different data set that used a color-word Stroop task from Gourley, Braver, and Bugg (2016) (Analysis 2).   Similar to Gonthier et al., the design revealed list-level control as PC 50 transfer items within a mostly congruent list had larger Stroop effects than PC 50 transfer items within a mostly incongruent list. For the second extension, I sought to determine if this method of quantifying the time scale of conflict learning could be applied to items within a list requiring item-level control (Analysis 3 and 4). The data used for Analysis 3 and 4 revealed item-level control by showing larger Stroop effects for mostly congruent compared to mostly incongruent items within the same list. The main prediction for item-level control is that the time scale of conflict learning ought to be short because it relies on reactive control processes. However, in addition, I sought to investigate whether items within the same list context may produce systematic differences in their time scales of control depending on the item's proportion congruence.

Analysis 1: Examining the List-Wide Time Scale of Conflict Learning in a Picture-Word Stroop Task from Gonthier et al. (2016)

Method

In Gonthier et al. (2016) each participant was exposed to three conditions. They comprised of a list-wide mostly congruent condition, a list-wide mostly incongruent condition, and a list-wide PC 50 condition. All three conditions included four PC 50 picture items. What differed across conditions was the PC of the other four picture items included in each list. In the list-wide mostly congruent condition, the other four were PC75. This led to an overall PC 67 list. In the list-wide mostly incongruent condition, the other four were PC 25.[1] This created an overall PC 33 list. The list-wide mostly congruent and incongruent conditions were comprised of 384 trials. For the third condition, the list-wide PC 50 condition, the other four were comprised of 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, this third condition was treated as a PC 50 list for comparison to the PC-67 and PC-33 lists. (Analysis 3 will analyze potential differences between the PC 75 and PC 25 items within the PC 50 list.) It bears mention, however, that this PC 50 list may not be representative of a typical PC 50 list in which every item is PC 50 within the list; in other words, the present PC 50 list may be more volatile than a typical PC 50 list wherein the same control setting can be applied to each item. The list-wide PC 50 condition was comprised of 432 trials. Additional information about the stimuli and design can be found in Gonthier et al.

---

[1] Note that behavioral patterns in Gonthier et al. (2016) indicated proportion congruence effects for PC 50 items in list-wide mostly congruent and list-wide mostly incongruent conditions. These patterns have been taken to indicate the operation of a list-level control mechanism as opposed to other mechanisms (e.g., contingency learning; see also Cohen-Shikora, Suh, & Bugg, 2018, for evidence countering a temporal learning mechanism).

Of the original 93 participants included in this dataset, 83 were retained for analysis. All 10 excluded participants were dropped because they did not complete all the conditions within the experimental paradigm. Additional demographic information can be found in Gonthier et al. (2016).

The following trials were excluded based on the exclusion criteria in Aben et al. (2017): the first trial of each condition (0.41%), error trials (3.77%), and trials following errors (3.77%). In addition, I box-trimmed the data such that trials <200ms and >3000ms were excluded (0.34%).

RTs are not typically normally distributed. Following the methods in Aben et al. (2017) I inverse transformed RTs (1/RT) and multiplied them by -10,000 to restrict the number of decimal places. This method has been suggested previously by Kinoshita, Mozer, &, Forester (2011) to better approximate normal distributions. With this transformation method, smaller inverse RTs reflect faster responses, so the direction of the scale is consistent with raw RTs.

First, collapsing across all conditions, I used a hierarchical linear model (HLM) to determine how many previous trials back I should use to model the timescale of control. The predictors in the HLM 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, I 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:

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

The results of this analysis are shown in Figure 3. To be more conservative, and departing from Aben and colleagues, instead of a one-tailed t-value, I used a two-tailed t-value of

1.96 for comparison. In the subsequent analysis I chose to include up to 16 trials back given the t-values were around that level until trial 16 (with a trial distance of 10 and 20 being exceptions).

Following this preliminary analysis, I statistically modeled RTs at two hierarchically related levels. The level one equation of the model was identical to that of the equation above with the caveat that k decreased from 24 to 16. This yielded the level one equation:

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

This equation 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, I was able to extract the coefficients of the interaction terms for each subject within each condition. These coefficients are the conflict adaptation weights (CAWs) I discussed earlier in this text. Each CAW at a given trial distance reflects the magnitude of shift in the congruency effect (in reaction time) on the present trial if an incongruent trial was presented at that trial distance.

Following Aben et al. (2017), at the second level, the 16 CAWs estimated by the level one equation were entered in as dependent variables in a 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. I was unable to implement a fully random structure 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. I was able to allow the effect of condition to be random, but I chose to only 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.

15

The effect of interest is the interaction between trial distance and condition. To assess the significance of this effect, I entered the predictors in the second level in a stepwise fashion and tested each model against its previous simpler nested model. Given the nested nature of the models, a test of the log likelihood ratio determines if the models are significantly different from one another with larger log likelihoods indicating more variance explained. Nonetheless, I also include the Akaike information criterion (Akaike, 1974) as a measure of model fit to mimic Aben et al. (2017). A smaller Akaike information criterion (AIC) indicates a better fit.

The Time Scale of Conflict Learning Results

The summary of the HLM model comparisons are shown in Table 1. The model with only log trial distance explained the variance in the data better than the model with just the intercept, $X^2(1) = 99.19$, $p < .001$. The model with only condition also explained the variance in the data better than the model with just the intercept, $X^2(2) = 22.46$, $p < .001$. The model with both main effects (log trial distance and condition) explained the data better than both the model with just log trial distance, $X^2(2) = 23.04$, $p < .001$, and the model with just condition, $X^2(1) = 99.77$, $p < .001$. Critically, there was a trending effect when comparing the full model (with the interaction) and the main effects only model, $X^2(2) = 5.40$, $p = .067$.

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) and Figure 5 displays the estimates without intercept correction for interested readers. I 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. This was accomplished using the glht() 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. There was a significant difference between the list-wide mostly congruent slope and list-wide mostly incongruent slope, $z = 2.98$, $p = .009$. There was no significant difference between the list-wide mostly congruent slope and list-wide PC 50 slope, $z = .97$, $p = .33$. There was a trending difference between the list-wide PC 50 slope and the list-wide mostly incongruent slope, $z = 2.01$, $p = .09$.

Discussion

The results from Analysis 1 are generally consistent with the findings from Aben et al. (2017). I observed a shorter time scale of conflict learning for the list-wide mostly congruent condition and a longer time scale of conflict learning for the list-wide mostly incongruent condition. This is made apparent by the more dramatic reduction in CAWs as trial distance increased in the list-wide mostly congruent condition compared to the list-wide mostly incongruent condition. This result indicates that the conflict learning mechanism places more importance or weight on recent events during low conflict situations, and it may change the relative weighting scheme to allow more distal events to play more of a role in high conflict situations.

The trajectory of the list-wide PC 50 condition suggests that this list had a time scale of control learning that was more similar to the list-wide mostly congruent condition than the list-wide mostly incongruent condition. When comparing the slopes of the trajectories, there was no difference between the list-wide PC 50 condition and list-wide mostly congruent condition, but there was a trend for the list-wide PC 50 condition to differ from the list-wide mostly

17

incongruent condition. Interestingly, the pattern of results for the list-wide PC 50 condition from Gonthier et al. (2016) appears to be similar to the volatile condition in Aben et al. (2017). Aben and colleagues posited that the volatile condition induced a greater reliance on recent trials (shorter time scale) due to its unstable nature. They suggested the volatile condition is unstable because the list switches back and forth between low conflict and high conflict conditions every 20 trials (see also, Jiang, et al., 2014). In unstable conditions, the cognitive control apparatus would have to shift quickly between substantially different control settings if it were to track context shifts. With that view, the time scale of control would benefit from being shorter in contexts where levels of conflict shift dramatically. Given this, the similarity of the pattern observed in the list-wide PC 50 condition in Gonthier et al. to the volatile condition in Aben et al. may not be surprising. In Gonthier et al., items within the list-wide PC 50 list were either PC 75 (mostly congruent), PC 50, or PC 25 (mostly incongruent), and Gonthier et al. reported evidence showing participants did use item-level control within this list. In other words, the control settings they used varied substantially from one item to the next within the list. Because the order in which items are presented is unpredictable, this condition can be thought of as volatile at the list-level.

The general replication of the findings of Aben et al. (2017), both in terms of showing differences in the time scales for mostly congruent and mostly incongruent lists and in showing that a novel type of volatile, PC 50 list behaved similarly to the volatile list in their study, provides evidence that the extended-CSE model is able to capture the time scale of conflict learning in tasks other than flanker. A potential caveat to note, however, is that the tasks used in Aben et al. (flanker) and 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 separated to

some degree, and thus may both involve spatial attention.. It could be argued that the extended-CSE model may not yield consistent patterns for qualitatively different tasks, such as those for which the dimensions are fully integrated (e.g., color word Stroop; Spieler, Balota, & Faust, 2000). However, if the model is capturing the time scale of conflict learning from a task-general control mechanism, then the model ought to result in similar patterns across cognitive control tasks. If the patterns seen here are not replicable across tasks, then the only way to salvage the interpretation that the extended-CSE model captures time scales is by adopting a non-parsimonious account of conflict learning that assumes task-specific learning mechanisms. Thus, it is important to investigate whether we can observe similar patterns when this model is applied to qualitatively different control tasks.

In Analysis 2, I applied the extended-CSE model to a dataset from Gourley et al. (2016) which used a color-word Stroop task. 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 integrated (as opposed to separated in the flanker and picture-word Stroop task) (MacLeod, 1998; Spieler, et al. 2000). If the pattern of results is similar in a color-word Stroop task, then the extended-CSE model can provide evidence that the time scale of control learning is part of a task-general control mechanism.

Analysis 2: Examining the List-Wide Time Scale of Conflict Learning in a Color-Word Stroop Task from Gourley et al. (2016)

Method

In Gourley et al. (2016) each participant was also exposed to three conditions. They comprised of a list-wide mostly congruent condition, a list-wide mostly incongruent condition, and a list-wide PC 50 condition. All three conditions included four PC 50 picture items. What differed across conditions was the PC of the other four picture items included in each list. In the list-wide mostly congruent condition, the other four were PC 75. 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 incongruent conditions were comprised of 288 trials. For the third condition, the list-wide PC 60 condition, the other four were comprised of two PC 100 items and two PC 25 items. The latter two sets of items were used to manipulate item-specific proportion congruence within this list. The list-wide PC 60 condition contained 480 trials.

All of the original 96 participants included in this dataset were retained for analysis. 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, I box-trimmed the data such that trials <200ms and >3000ms were excluded (0.82%). The RTs were transformed in the same fashion as in Analysis 1.

As in Analysis 1, I first sought to determine how many previous trials back I should use for modeling the time scale of control. Again, I chose to look at up to 24 previous trials back as an initial starting point. To keep the analysis consistent with Analysis I, the HLM was

implemented with only the intercept being allowed to vary on a subject level. This resulted in the level one equation:

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

The results of this analysis are shown in Figure 6. For the time scale of conflict learning analysis, I chose to include up to 8 trials back given the t-values were above 1.96 through trial 8.

Following this preliminary analysis, for the time scale of conflict learning, I statistically modeled RTs at two hierarchically related levels. The level one equation of the model was identical to that of the equation above with the caveat that k decreased from 24 to 8. This yielded the level one equation:

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

This equation 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, I was able to extract the coefficients of the interaction terms for each subject within each condition.

At the second level, the 8 CAWs estimated by the level one equation were entered in as dependent variables in a 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. I chose to only allow the intercept to be random to keep the analysis procedure consistent with Analysis I and Aben et al. (2017). To assess the significance of the log trial distance x condition interaction, I entered the predictors in the second level in a stepwise fashion and tested each model against its previous simpler nested model.

The Time Scale of Conflict Learning Results

The summary of the HLM model comparisons are shown in Table 3. The model with only log trial distance explained the variance in the data better than the model with just the intercept, $X^2(1) = 127.48$, $p < .001$. The model with only condition also explained the variance in the data better than the model with just the intercept, $X^2(2) = 12.44$, $p = .002$. The model with both main effects of log trial distance and condition explains the data better than both the model with just log trial distance, $X^2(2) = 13.18$, $p = .001$, and the model with just condition, $X^2(1) = 128.22$, $p < .001$. Importantly, there was a significant effect when comparing the full model (with the interaction) and the main effects only model, $X^2(2) = 19.74$, $p < .001$.

The regression coefficients of the full model are displayed in Table 4. Figure 7. displays the estimates of the full model corrected for the intercept of each condition and Figure 8. displays the estimates without intercept correction. Using a no intercept model and the glht() function I investigated the pairwise differences in slopes. There was a significant difference between the list-wide mostly congruent slope and list-wide mostly incongruent slope, z = 2.66, *p* = .02. There was also a significant difference between the list-wide PC 60 slope and the list-wide mostly incongruent slope, z = 4.73, *p* < .001. Interestingly, there was also a significant difference between the list-wide mostly congruent slope and the list-wide PC 60 slope, z = 2.08, *p* = .04.

Discussion

I observed a shorter time scale of conflict learning in the list-wide mostly congruent condition compared to the list-wide mostly incongruent condition. I also observed a shorter time scale of conflict learning in the list-wide PC 60 condition compared to the list-wide mostly

22

incongruent condition. These results align with the results from Analysis 1 and Aben et al. (2017). Interestingly, differing from the PC 50 list in Analysis 1, the PC 60 list had a shorter time scale than the list-wide mostly congruent condition. This may seem surprising since the list-wide mostly congruent condition is PC 67. That is, the list-wide mostly congruent condition is less conflicting than the list-wide PC 60 condition. Thus far, the patterns from this model have indicated that a decrease in conflict levels shortens the time scale of conflict learning reflecting shifts away from proactive control and towards reactive control. However, the list-wide PC 60 condition is also an unstable condition because it biases use of item-level control (given the inclusion of items with different proportion congruencies in the list). With the perspective that there is an additive effect of list-level conflict and list-level instability for shifts to reactive control, then it could be that the list-level instability of the PC 60 list shortened the time scale of conflict learning enough to compensate for and "overcome" the .07 proportion congruence difference between the PC 67 list-wide mostly congruent list and the list-wide PC 60 list. Alternatively, given that this was not observed in Analysis 1, it could be a Type 1 error.

Given the general concordance of results for list-wide mostly congruent and incongruent conditions from Analysis 1, 2, and Aben et al. (2016), the extended-CSE model appears to be robust enough to capture time scale patterns that ought to be similar across qualitatively different cognitive control tasks if there is a task-general control mechanism. Still, a lingering question that remains is do time scale differences emerge for *items* that have different levels of conflict within the same list? The motivation for assuming different time scales for list-level control emerges from behavioral differences that indicate that different attention settings are applied to mostly congruent and mostly incongruent lists. As mentioned earlier, similar behavioral differences are present when comparing mostly congruent and mostly incongruent *items* within a

list which indicate item-level control. However, 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 the item in the current trial (see Figure 9). Only the conflict statuses of previous instances of that item would be informative. Consequently, if there are differences in the time scale of conflict learning for items with different proportion congruences, then differences in the CAW slopes would be seen in previous *instances* of an item and not previous trials. To investigate time scale differences for items with differing proportion congruences, I applied the extended-CSE model selectively to the list-wide PC 50 condition from Gonthier et al. (2016) which was designed to investigate item-level control. In addition, I modified the model to use log *instance* distance (as opposed to log trial distance) as a predictor. I could not apply the extended-CSE model to the Gourley et al. (2016) list-wide PC 60 condition, which also was designed to investigate item-level control, because mostly congruent items in that condition were PC 100 and thus there were no incongruent trials to model for those items.

Analysis 3: Examining the Item-Specific Time Scale of Conflict Learning in a Picture-Word Stroop Task from Gonthier et al. (2016)

Method

As described in Analysis 1, the PC 50 list in the Gonthier et al. (2016) study included an item-specific proportion congruence manipulation whereby two picture items were mostly congruent and two picture items were mostly incongruent[2]. As also described previously, there were four PC 50 items; however, given that there were only 12 instances of a particular PC 50 item in the list (e.g., only 12 instances of the color *blue*), I excluded PC 50 items from this analysis and only examined mostly congruent and mostly incongruent items. 89 participants who completed the item-specific proportion congruence (PC 50 list) condition were included in this analysis. The following trials were excluded: the first trial of each condition (0.16%), error trials (3.44%), and trials following errors (3.44%). In addition, I box-trimmed the data such that trials <200ms and >3000ms were excluded (0.32%). The RTs were transformed in the same fashion as in Analysis I and II.

There are a fourth the number of instances of an item in the item-specific proportion congruence condition as there are total trials in the list-wide mostly congruent condition. Therefore, it did not seem reasonable to try and determine how many instances back I should look using t-value cutoffs as I did in Analysis 1 and 2. Thus, I chose to look back up to 6

---

[2] The list-wide PC 50 condition in Gonthier et al. (2016) was designed such that the relevant dimension of items (picture) differentially predicted proportion congruence (whether an item was mostly congruent or mostly incongruent) but it did not differentially predict responses. That is, the pictures were correlated with the correct response 100% of the time, regardless of whether the item was mostly congruent or mostly incongruent (for further explication, see Bugg, 2012; Bugg et al., 2011). As detailed in the dual item-specific mechanism account of item-specific proportion congruence effects (Bugg, 2015; Bugg & Hutchison, 2013; Bugg et al., 2011), this type of design produces a control-based item-specific proportion congruence effects and not a contingency-based effect (see also Chiu et al. 2017).

instances (a fourth of my original maximum for list-wide analyses) for the time scale of conflict learning analysis. The resulting level one linear model for the time scale analysis was:

$$RT = \beta_0 + \beta_1 C_i + \beta_2 C_{i-1} + \beta_3 C_{i-2} \dots + \beta_7 C_{i-6} + \beta_8 C_i C_{i-1} + \beta_9 C_i C_{i-2} \dots + \beta_{13} C_i C_{i-6}$$

Here, 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). This equation was applied iteratively to each subject for each item type (mostly congruent and mostly incongruent). In doing so, I was able to extract the coefficients of the interaction terms for each subject for each item type.

At the second level, the 6 CAWs estimated by the level one equation were entered in as dependent variables in a HLM with log instance distance and item type as predictors. I entered the predictors in the second level in a stepwise fashion and tested each model against its previous simpler nested model.

The Time Scale of Conflict Learning Results

The summary of the HLM model comparisons are shown in Table 5. The model with only log instance distance explained the variance in the data better than the model with just the intercept, $X^2(1) = 5.63$, $p = .018$. The model with only item type showed trending evidence of explaining the variance in the data better than the model with just the intercept, $X^2(1) = 3.69$, $p = .054$. The model with both main effects of log instance distance and item type also showed trending evidence that it explained the data better than the model with just log instance distance, $X^2(1) = 13.18$, $p = .001$. The model with both main effects did explain the data better than the model with just item type, $X^2(1) = 5.64$, $p = .017$. Importantly, there was no significant difference when comparing the full model (with the interaction) and the main effects only model, $X^2(1) = .48$, $p = .488$.

The regression coefficients of the full model are displayed in Table 6. Figure 10 displays the estimates of the full model corrected for the intercept of each condition and Figure 11. displays the estimates without intercept correction. The t-value for the interaction shows no significant difference between the slopes of the item types.

Discussion

I observed no difference in the time scale of conflict learning for mostly congruent and mostly incongruent items within a list. One interpretation of this result is that, for items within a list that bias adoption of item-level control, the time scale of conflict learning does not differ as a function of probability of conflict for specific items. Note that this account does not exclude the possibility that the time scale of conflict learning for items is dependent on the global proportion congruence of the list in which the items reside (as demonstrated in Analysis 1). An equally plausible interpretation is that the extended-CSE model does not capture the time scale for items. This account would imply that conflict learning differences are not evident from the independent influence of previous instances and another model would better capture the time scale for items within a list requiring item-level control.

Another possibility is that differences in the time scale could emerge with this modeling procedure, but they are not observable in the Gonthier et al. (2016) data because of insufficient power. The results indicated no significant difference for mostly congruent items and mostly incongruent items, but the direction of the pattern was visually similar to the list-wide time scale pattern (see Figure 4 and 7). To address this possibility, in Analysis 4, I examined data from Bugg and Dey (2018) which also used an item-specific proportion congruence manipulation

within a picture-word Stroop task but with many more participants to determine if the results from Analysis 3 were replicable.

Analysis 4: Examining the Item-Specific Time Scale of Conflict Learning in a Picture-Word Stroop Task from Bugg and Dey (2018).

Method

I merged data from five experiments in Bugg and Dey (2018, Exp 1, Exp 2, Exp 3b, Exp 4b, and Exp 5). The Stroop phase of these experiments was identical and the behavioral results in RTs yielded the same effects across experiments, and as in Gonthier et al. (2016). Specifically, there was 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. There were four picture items (bird, cat, dog, and fish). These items were split into two sets (bird-cat, dog-fish). One item set was mostly congruent and the other set was mostly incongruent. The mostly congruent items were PC 67 and the mostly incongruent items were PC 33 such that the overall list was PC 50, as in Gonthier et al. (2016; Analysis 3).[3] There were no PC 50 items in these experiments.[4] Participants were presented with 432 trials in total. Additional details on the design and stimuli can be found in Bugg and Dey (2018).

Aggregating across the five experiments, there were 216 participants who were included in this analysis. 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, I box-trimmed the data such

---

[3] Similar to the list-wide PC 50 condition in Gonthier et al. (2016), the item-specific proportion congruence designs used in Bugg and Dey (2018) produced control-based item-specific proportion congruence effects.
[4] In Bugg and Dey (2018) picture stimuli could be treated as members of a category of pictures (e.g., bird, cat, dog, or fish) or individual exemplars within a category (e.g., robin, oriole, or pigeon). PC 50 exemplars were presented in this experiment, but they were always members of either a mostly congruent or mostly incongruent category. Here we treat the category as the item.

that trials <200ms and >3000ms were excluded (0.39%). The RTs were transformed in the same

fashion as in Analysis 1, 2, and 3. To keep consistent with Analysis 3, I chose to look back up to

6 instances.

The Time Scale of Conflict Learning Results

The summary of the HLM model comparisons are shown in Table 7. The model with

only log instance distance explained the variance in the data better than the model with just the

intercept, $X^2(1) = 35.94$, $p < .001$. There was no difference between the model with only item

type and the model with just the intercept, $X^2(1) = .24$, $p = .618$. There was no difference

between the model with both main effects of log instance distance and item type and the model

with just log instance distance, $X^2(1) = .25$, $p = .615$. The model with both main effects did

explain the data significantly better than the model with just item type, $X^2(1) = 35.94$, $p < .001$.

Importantly, as in Analysis 3, there was no significant difference when comparing the full model

(with the interaction) and the main effects only model, $X^2(1) = .41$, $p = .522$.

The regression coefficients of the full model are displayed in Table 8. Figure 12 displays

the estimates of the full model corrected for the intercept of each condition and Figure 13

displays the estimates without intercept correction. The t-value for the interaction shows no

significant difference between the slopes of the item types.

Discussion

I observed no difference in the time scale of conflict learning for mostly congruent and

mostly incongruent items which replicated the results from Analysis 3. This provides evidence

that the null result from Analysis 3 is unlikely to be attributable to insufficient power. However,

the results from Analysis 4 do not adjudicate between the two interpretations that were proposed

for why there is no difference in the time scale of conflict learning for mostly congruent and mostly incongruent items. One interpretation is that the time scale of conflict learning is not modulated by the probability of conflict for specific items, while another interpretation is that the extended-CSE model is an inappropriate method for capturing time scale differences at the item-level. I will consider these two interpretations further in the General Discussion.

General Discussion

The set of analyses presented in this paper aimed to accomplish three goals. First, I attempted to replicate time scale patterns for list-wide proportion congruence manipulations observed in Aben et al. (2017) using their extended-CSE model. Second, I attempted to validate the model across different tasks to provide evidence that the model indexed a task-general conflict learning mechanism. Third, I extended their model to attempt to describe the time scale of conflict learning for items within a list that required item-level control. Combined, these goals served to further refine extant theoretical and computational models of a general cognitive control apparatus. Specifically, I attempted to better characterize the conflict learning mechanism that is assumed to be required for the control system to adjust attention settings based on statistical regularities of conflict.

Differences in the Time Scale of Conflict Learning between List-Wide Mostly Congruent and List-Wide Mostly Incongruent Contexts

Analyses 1 and 2 converged on the notion that low conflict contexts have a short time scale for conflict learning as compared to high conflict contexts. This was evident in differences in the slopes of the CAW trajectories for list-wide mostly congruent and list-wide mostly

30

incongruent conditions, with the list-wide mostly congruent conditions showing more dramatic reductions in CAW as trial distance increased. Importantly, this pattern was evident across two qualitatively different tasks (picture-word Stroop and color-word Stroop). This provides evidence that the extended-CSE model (Aben et al., 2017), a model that was validated initially with a flanker task, captures the time scale of conflict learning at the list-level regardless of the nature of the conflicting stimuli (i.e., from fully separable to fully integrated dimensions). This suggests the model may be indexing task-general control mechanisms.

An important assumption for this interpretation (i.e., that the global probability of conflict is related to different time scales of learning) is that the relative reduction in CAWs across trial distance (i.e., the slopes of the trajectories across trials) is the measure that captures time scale. This measure ignores the relative difference in the magnitude of CAWs averaging across trial distances (i.e., the main effect of condition). Figures 5 and 8 plot the CAWs of the list-wide conditions that include the main effect of condition. These figures show that within a window size of 16 and 8, respectively, all previous trials in the list-wide mostly congruent condition showed a nominally larger independent effect of incongruent trials compared to the independent effect of incongruent trials in the list-wide mostly incongruent condition. Within these window sizes, the trajectories did not cross such that the independent effect at a distal trial distance was nominally less in the list-wide mostly congruent condition. The interpretation I forward here is that the main effect of condition reflects the difference in overall levels of attention, whereas the change in CAWs, or the slope, is more reflective of learning. To reiterate, a CAW at a trial distance of k is the independent effect of the kth previous trial if that trial were incongruent. More negative average CAWs would thus mean that larger adjustments are made following an incongruent trial. In the list-wide mostly congruent condition, incongruent trials are infrequent

and thus a relaxed attentional state would usually be preferred to minimize taxing cognitive resources. Therefore, the presentation of an infrequent incongruent trial would "jump-start" the control system to produce a reflexive/reactive increase in attention. This would be observed as more negative CAW values. In contrast, in the list-wide mostly incongruent condition, incongruent trials are frequent and thus a focused attentional state would be preferred to maintain accuracy.   In this condition, less negative average CAW values would be predicted because there is less need for adjustments to be made during a focused state. Thus, the difference in the average magnitude of CAWs (the main effect of condition) indexes the difference in average levels of attention (within a given window size), with less negative CAWs reflecting a more focused state. In contrast, steeper shifts in CAWs for from trial-to-trial for the list-wide mostly congruent condition reflect greater weighting of recent trials *relative* to distal trials. If the conflict learning mechanism accumulates conflict information using an exponentially decaying function, as assumed by reinforcement learning models, this would mean that the learning rate is set such that control adjustments are more reliant on the first few previous trials, and less reliant on the global probability of conflict within the condition. Comparatively shallower shifts in CAWs in the list-wide mostly incongruent condition reflect relatively less reliance on recent trials and more reliance on the global probability of conflict within the condition.

Time Scale of Conflict Learning for "Volatile" Lists that Bias Adoption of Item-Level Control

Analysis 1 and 2 revealed that list-wide PC 50 and list-wide PC 60 lists, which had list-level proportion congruences in-between mostly congruent and mostly incongruent lists, produced time scales of conflict learning that were more similar to time scales produced by list-wide mostly congruent lists than those produced by list-wide mostly incongruent lists. This may

32

be an unsurprising result if one takes the view that time scales are predominantly impacted by the type of control that is engaged and less so by the probability of conflict per se. The list-wide PC 50 and list-wide PC 60 lists were composed of items with very different item-specific proportion congruences (i.e., either PC 75 or PC 100 for mostly congruent items and PC 25 for mostly incongruent items). Similar to list-wide mostly congruent conditions, lists with very different item-specific proportion congruences are thought to engage more reactive control processes. Thus, a strong interpretation of the results from the present research would be that more engagement of reactive control processes leads to a short time scale and more engagement of proactive control processes leads to a long time scale regardless of the probability of conflict experienced. For example, an individual relying more heavily on proactive control processes in a list-wide mostly congruent condition may express a long time scale of conflict learning, which would contradict the group-level patterns for list-wide mostly congruent conditions found in the present research.

Time Scale of Conflict Learning for Mostly Congruent and Mostly Incongruent Items within a List

In Analysis 3, I observed no difference in the time scales for items with vastly different probabilities of conflict within the same list. That is, mostly congruent items expressed the same time scale as mostly incongruent items within a list. Both types of items are thought to trigger adjustments to attention reactively and not proactively. That I did not observe any differences is consistent with the view that the time scale of conflict learning is contingent on the type of control engaged and not the probability of conflict. This interpretation is unlikely to be founded on a Type II error as the results were replicated in Analysis 4 using a larger data set. The observed time scale invariance demonstrates that differential modulations of attention for *items*

are not influenced by different learning rates for items. This is consistent with an episodic

retrieval account of cognitive control, forwarded by Crump and Milliken (2009), which would

assume that items become bound to attentional settings over time, and the rate at which the

association between item and attentional setting is formed does not differ based on the item type

(mostly congruent or mostly incongruent). Additionally, these results may point to separable yet

concurrent conflict learning processes for the list-level and the item-level. That is, Analysis 1

and Analysis 3 combined demonstrate that the congruency effect of current trials is impacted by

both previous trials *and* previous instances (although these analyses do not demonstrate

independent effects).   Further research is needed, however, to strengthen these conclusions as

the above interpretations rely on the assumption that the modification I made to the extended-

CSE model is an appropriate method for quantifying time scales of conflict learning for

individual items. I have only used two data sets to determine this method's validity in capturing

time scales of conflict learning at the item-level, and so my conclusions may be tentative until

this method is applied to additional lists from experiments that bias adoption of item-level

control.

Implications for Extant Models of Cognitive Control

        The results of this present research have important implications for many conflict-

triggered cognitive control models. The conflict-monitoring hypothesis (Botvinick et al. 2001) is

a seminal account of cognitive control because it allows for the quantification of conflict and

demonstrates that conflict monitoring explains behavioral and neuroimaging data for list-level

control processes. The conflict-monitoring hypothesis assumes that there is a monitoring unit

that tracks conflict and appropriately adjusts attention at the pathway level (e.g., color or word

pathway in color-word Stroop). The item-specific adaptation of the conflict monitoring hypothesis (Blais et al. 2007) modified the original model to also account for item-level control processes by assuming that attention adjustments are made at the level of items (e.g., word reading is attenuated for mostly congruent items). A conflict-modulated Hebbian learning model (Verguts & Notebaert, 2008) has also been forwarded and accounts for global pathway-level control (different tasks in a task switching paradigm) and item-level control.

Importantly, and relevant for the present research, none of the above models specified how many trials (or instances) back the conflict monitoring unit utilizes to determine levels of conflict (or probability of conflict). The present results concur with the results from Aben et al. (2017) in demonstrating the ability to quantify time scales of conflict learning and by showing that time scale parameters systematically vary based on the type of control engaged. These findings provide a basis for updating extant conflict-modulated accounts of cognitive control. Specifically, the results indicate that using a single learning rate parameter for a list-level conflict monitoring unit is not optimal because low conflict lists produce short time scales and high conflict lists produce long time scales. In other words, low conflict contexts seem to produce a large learning rate, such that the level of conflict is determined by relatively recent trials, and high conflict contexts seem to produce a small learning rate such that the level of conflict is determined by both recent and distal trials. This is consistent with the dual mechanisms of control account (De Pisapia & Braver, 2006) which predicts long time scales for proactive control processes and short time scales for reactive control processes. In addition, according to the item-level time scale results (Analyses 3 and 4), the pattern of conflict learning for items does not follow the pattern of list-level conflict learning. That is, there are no differences in time scales based on items with differing probabilities of conflict. Thus, it is possible that there are

35

two conflict monitoring units, one at the list-level (or at the global pathway level), and one at the item-level. Following this interpretation, when there are high levels of conflict, the analyses in the present study indicate that a list-level conflict monitoring unit and an item-level conflict monitoring unit use differing amounts of information. That is, when calculating the probability of conflict for a given trial, a list-level conflict monitoring unit uses both recent and distal trials in lists with a high probability of conflict. In comparison, when calculating the probability of conflict for a given instance, an item-level conflict monitoring unit uses mostly recent instances for an item with a high probability of conflict. This may be unsurprising if one takes the view that information about conflict on previous instances or trials degrades over time. Since previous instances of an item are more spread out temporally than previous trials in a list, a conflict monitoring unit may only use recent instances by necessity because conflict information from more distal instances have been lost in some sense. Future replications of this work notwithstanding, these contrasting findings for list-level and item-level control further support that these two mechanisms are dissociable (cf. Gonthier et al., 2016).

Limitations and Future Directions

The conclusions of the present research regarding the time scale of conflict learning, as well as those of Aben et al. (2017), hinge on the notion that the CSE is a behavioral marker for adjustments of attention. In other words, the CSE reflects the output of a cognitive control mechanism (Blais et al., 2007; Botvinick et al., 2001; Verguts & Notebaert, 2008). Other accounts posit that the CSE is more reflective of simple stimulus-response processes and control need not be invoked to explain observed effects (Hommel, Proctor, & Vu, 2004; Mayr, Awh, & Laurey, 2003; Schmidt & De Houwer, 2011). One such account is the feature-integration

account (Hommel et al., 2004) which posits that the features of the stimulus and response of a trial get bound together in an event file. In the next trial, when all the features (complete repetition) or none of the features (complete alternation) match the event file, processing is facilitated relative to trials where there is some overlap (partial repetition) with the event file. In a two-alternative-forced choice task (e.g., a two-choice color-word Stroop task), congruent-congruent sequences and incongruent-incongruent sequences will always be either complete repetitions or complete alternations. In contrast, congruent-incongruent and incongruent-congruent sequences will always be partial repetitions. Thus, the CSE can be explained by processes that do not require a control mechanism in two-alternative-forced choice tasks. However, when moving to four-alternative-forced choice tasks, it has been demonstrated that CSE is still observed either by preventing feature repetitions from occurring in sequence (Jiménez & Méndez, 2013) or by removing feature repetitions from sequential analyses (Akcay & Hazeltine, 2007).

But, the CSE in four-alternative-forced choice tasks can be explained by another associative account – the contingency account (Schimidt & De Houwer, 2011). The contingency account claims that the irrelevant stimulus feature (e.g., the word BLUE) becomes associated with its congruent response (i.e., the color blue) because there are more presentations of *each* congruent stimulus than there are of *each* incongruent stimulus (for review, see Egner 2017). For example, in the list-wide mostly congruent condition from Gonthier et al. (2016) the word CAT was presented with a cat picture 54 times and was presented with a dog, fish, and bird picture six times each. Thus, instead of differentiating stimuli based on conflict (i.e., congruent or incongruent stimuli), one can differentiate stimuli based on contingency (i.e., high-contingency or low-contingency stimuli). One can then re-express the CSE as a *contingency*

sequence effect as it has been shown that high-contingency trials are responded to faster than low-contingency trials, and contingency constant sequences (i.e., low-contingency – low-contingency, or high-contingency – high-contingency) facilitate performance (Schimidt & De Houwer, 2011). A way to demonstrate the CSE controlling for contingency and feature repetitions is by using a four-alternative-choice task that is comprised of two non-overlapping two-alternative-choice sets that alternate in sequence (Schmidt & Weissman, 2014). Importantly, in both the present research and in Aben et al. (2017), a design controlling for both a feature-integration account and a contingency account of the CSE was not used. That is, on some subset of trial sequences the feature-integration account could explain the CSE due to partial repetitions, and for some lists and items, contingency learning could be driving the CSE because words co-occurred with congruent responses (picture or color) more often than other incongruent responses.

However, all the datasets in the present research and the data from Aben et al. (2017) used list-wide or item-specific manipulations that produced list-wide and item-specific effects that cannot be accounted for by contingency learning. That is, in the list-wide manipulations, transfer of proportion congruence effects to PC 50 items was observed which rules out a contingency account of list-wide proportion congruence effects. Likewise, in the item-specific manipulations, the co-occurrence of words with their respective congruent responses did not allow for contingency learning, ruling out a contingency account of item-specific proportion congruence effects. Thus, to claim that contingency learning is influencing the time courses of CAWs observed here, one would have to assume contingency learning influences CAWs even though it cannot account for overall behavior (i.e., list-wide proportion congruence and item-specific effects).

38

The extended-CSE model is more powerful than conventional CSE and proportion congruence models as it uses more of the data to estimate the parameters of interest. Specifically, it uses trial level data and includes more than a single previous trial when predicting behavior of a current trial, whereas conventional CSE models use only a single trial back and proportion congruence models collapse behavior across many trials. Because of this, it is fine-grained since we can obtain CAWs for any number of previous trials back. However, this also comes with limitations. The number of parameters that are estimated by the level one equation in the model is quite large. As a result, the CAWs extracted for each person and each condition must be generated by individual ordinary least-squares regressions. These coefficients are not adjusted for by the reliability of observations and are more susceptible to outliers (Snijders & Bosker, 2012). In comparison, in a true 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 true hierarchical linear modeling procedure cannot be applied to the extended-CSE model because it requires many more trials per subject than what is practically feasible. Hierarchical linear models fail to converge when there are not enough observations for the number of specified parameters to be estimated. In Aben et al. (2017) the number of parameters that were estimated for each participant in each condition was 26. In Analysis 1 and 2, the number of parameters to be estimated was 34 and 18, respectively. In Analysis 3 and 4, the number of parameters to be estimated was 14. In all analyses presented here, the model failed to converge when applying a true hierarchical linear modeling procedure.

Alternative models aiming to characterize the time scale of conflict learning would benefit from reducing the number of parameters to be estimated while taking advantage of as much of the data as the extended-CSE model does. Examples of potential models include those

that use learning rate parameters as a means of indexing the time scale of control (Chiu, Jiang, &

Egner, 2017; Jiang, Beck, Heller, & Egner, 2015; Jiang, et al., 2014). Drawing inspiration from

these models, one could assume that the proportion congruence estimated for the current trial

(which would inform selection of an attentional state) is a function of a weighted combination of

the previous trial and the previous estimated proportion congruence. The weighting is

determined by a learning rate. For example, the function may be formalized as,

$lPC_i = \lambda C_i + (1-\lambda)lPC_{i-1}$, where, lPC_i is the local proportion congruence as of the current trial. C_i

is the congruency state of the current trial. And, lPC_{i-1} is the local proportion congruence as of

the last trial. Here, λ serves as the learning rate such that when λ is large the local proportion

congruence as of the current trial is mostly determined by more recent trials. Essentially, λ

parameterizes the time scale of control. 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.

Appropriately adjusted and more accurate estimates may then be extracted from the same

number of observations and individual differences could be extracted with a large enough sample

size.

Conclusion

In conclusion, I applied an extended-CSE model developed by Aben and colleagues

(2017) to two qualitatively different cognitive control tasks and demonstrated the model's ability

to consistently capture the time scale of conflict learning. I replicated the results from Aben and

colleagues and observed a) a relatively short time scale for low conflict list-wide contexts (i.e.,

mostly congruent lists) and volatile list-wide contexts, and b) a relatively long time scale for high

conflict list-wide contexts (i.e., mostly incongruent lists). I discussed how these findings point to

a task-general conflict learning mechanism. I also applied the extended-CSE model to mostly

congruent and mostly incongruent items within a list and found no differences in the time scale

of items with vastly different probabilities of conflict. Based on the collective pattern of results, I

forwarded the idea that the conflict learning mechanism modulates time scales of conflict

learning based on whether proactive or reactive control processes are recruited but does not

modulate the time scale based on the nature of the reactive control process that is recruited (i.e., a

focused vs. relaxed attentional setting). This model provides a means of quantifying conflict

learning which can subsequently be used to refine extant mechanistic and theoretical models of

cognitive control by empirically determining learning rate parameters for conflict monitoring

units.

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.

Akcay, C., & Hazeltine, E. (2007). Conflict monitoring and feature overlap: Two sources of

sequential modulations. *Psychonomic Bulletin & Review, 14*, 742-748.

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

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.

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.

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. (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, UK: 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., & 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. (2011). 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. (2011). 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.

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.

Crump, M. J. C., Gong, Z., & Milliken, B. (2006). The context-specific proportion congruent Stroop effect: Location as a contextual cue. *Psychonomic Bulletin & Review, 13*, 316 – 321.

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

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.

Cohen, J. D. (2017). Conflict adaptation: Past, present, and future of the congruency sequence

    effect as an index of cognitive control. In T. Egner (Ed.), *The Wiley handbook of*

    cognitive control. (pp. 64-78). Chichester, West Sussex, UK: John Wiley & Sons.

Egner, T., & Hirsch, J. (2005). Cognitive control mechanisms resolve conflict through cortical

    amplification of task-relevant information. *Nature Neuroscience, 8,* 1784–1790.

Gonthier, C., Braver, T. S., & Bugg, J. M. (2016). Dissociating proactive and reactive

    control in the Stroop task. *Memory & Cognition, 44*, 778 − 788.

Gourley, E. M., Braver, T. S., & Bugg, J. M. (2016, November). 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.

Hothorn, T., Bretz, F., & Westfall, P. (2008). Simultaneous inference in general parametric

    models. *Biometrical Journal, 50,* 346-363.

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.

Kane, M. J., & Engle, R. W. (2003). Working-memory capacity and the control of attention: The contributions of goal neglect, 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.

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.

Mayr, U., Awh, E., & Laurey, P. (2003). Conflict adaptation effects in the absence of executive

    control. *Nature Neuroscience, 6,* 450-452.

Melara, R. D., & Algom, D. (2003). Driven by information: A tectonic theory of Stroop

    effects. *Psychological Review, 110*, 422-471.

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.

Schmidt, J. R., De Houwer, J. (2011). Now you see it, now you don't: Controlling for

    contingencies and stimulus repetitions eliminates the Gratton effect. *Acta Psychologica,*

    138, 176–186.

Snijders, T. A., & Bosker, R. J. (2012). *Multilevel analysis: An introduction to basic and*

    advanced multilevel modeling. Los Angeles: SAGE.

Tzelgov, J., Henik, A., & Berger, J. (1992). Controlling Stroop effects by manipulating expectations for color words. *Memory & Cognition, 20,* 727–735.

Verguts, T., & Notebaert, W. (2008). Hebbian learning of cognitive control: Dealing with specific and nonspecific adaptation. *Psychological Review, 115*, 518-525.
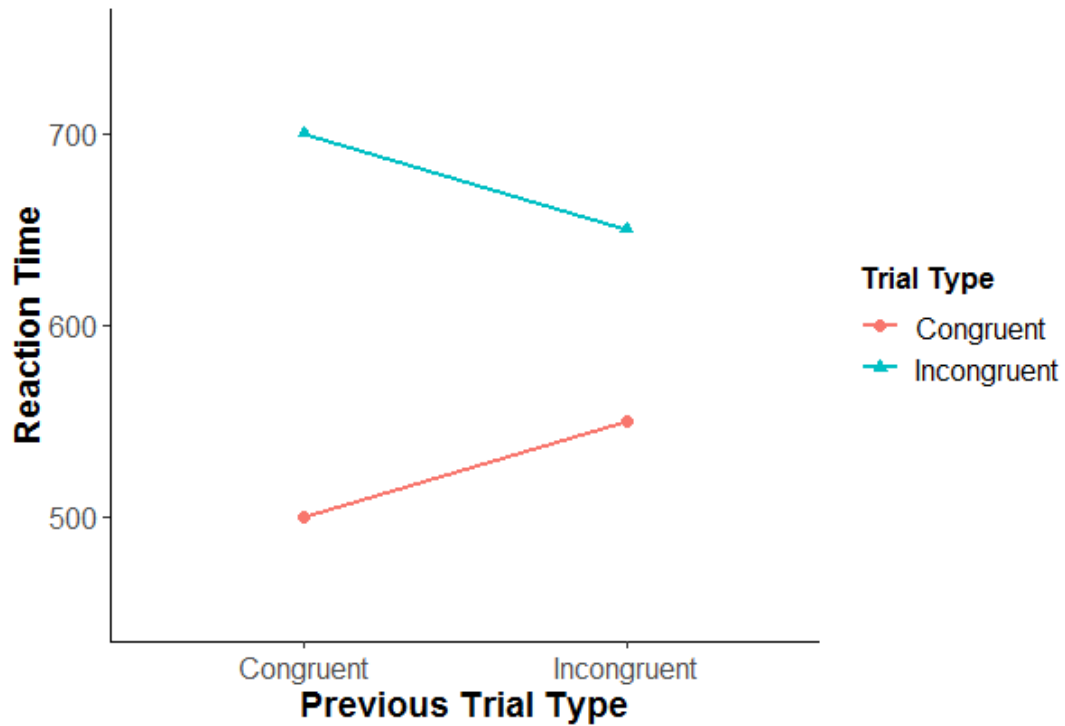
Figure 1. Hypothetical data depicting the congruency sequence effect (CSE). Plotted on the x axis is the trial type of the preceding trial. Red circles depict current congruent trials. Blue triangles depict current incongruent trials. The difference in red and blue points for each previous trial type 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.

Figure 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 change in CAW across trial distance (i.e., the slopes of the trajectories) index the time scale of conflict learning. Long time scales are indexed by relatively smaller changes across trial distance. Short time scales are indexed by relatively steeper changes across trial distance.

Figure 3. C_i C_{i-k} interaction t values of the level one model (*k* ranges from 1 - 24), plotted as a

function of trial distance. 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
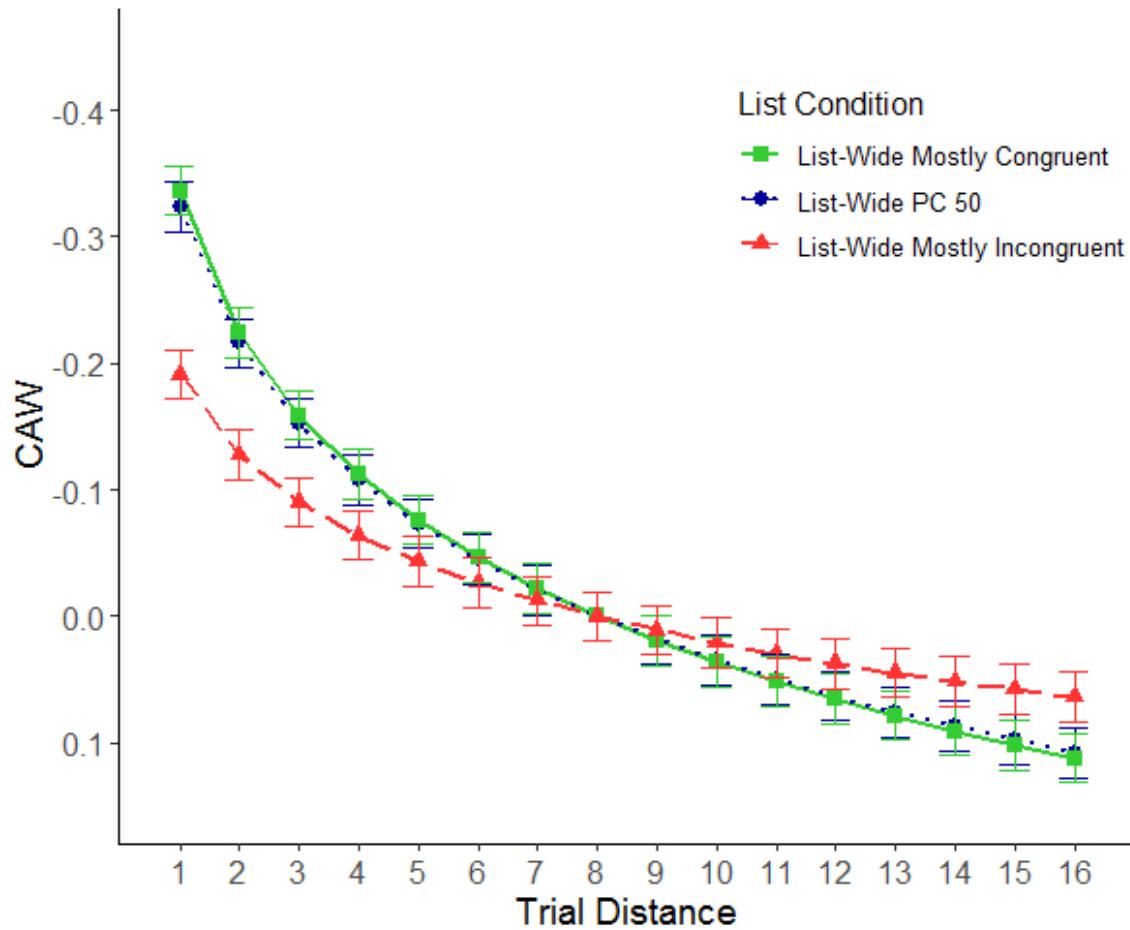
Gonthier et al. (2016).

Figure 4. Estimates of the CAWs in Model 4 (full model) using data from 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 linear mixed models for
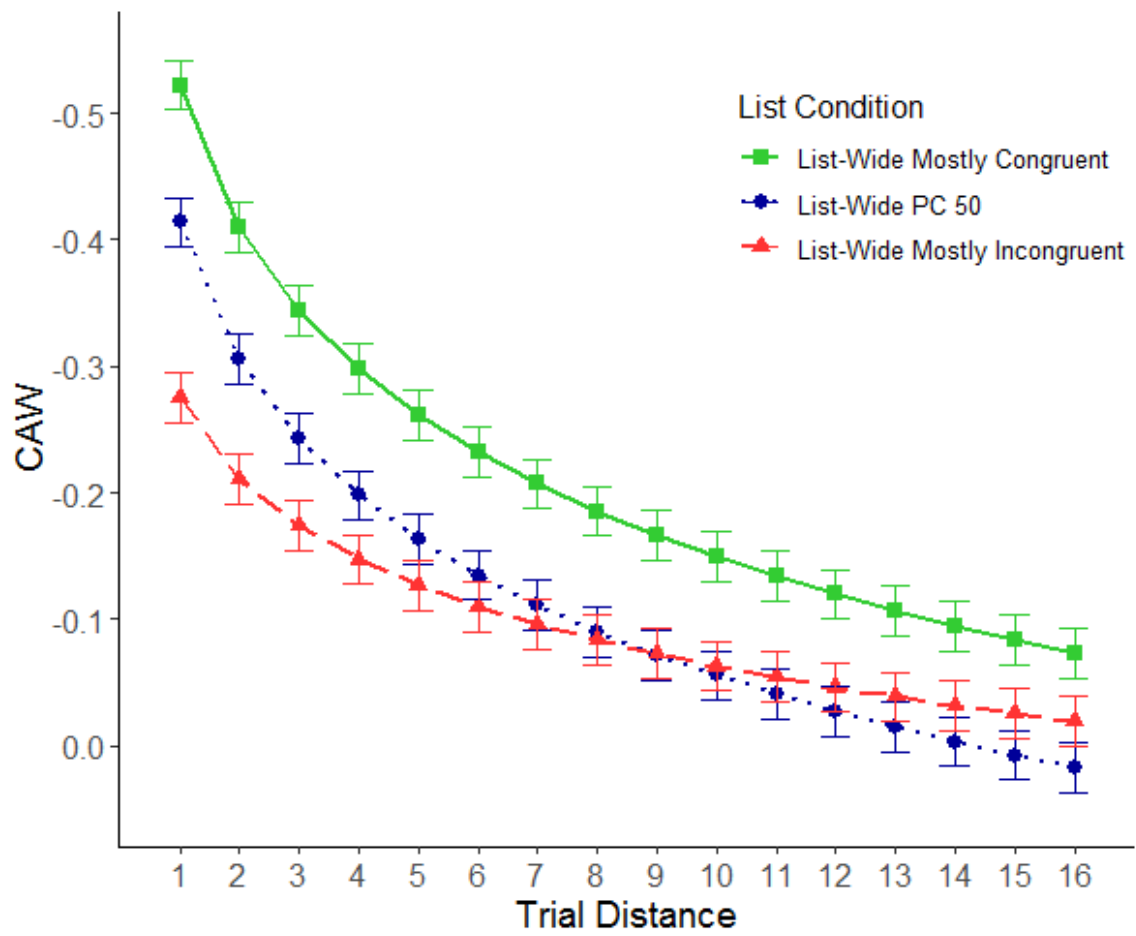
statistical testing).

Figure 5. Estimates of the CAWs in Model 4 (full model) using data from Gonthier et al. (2016).

No intercept correction was applied. The original scale of trial distance is displayed on the x-
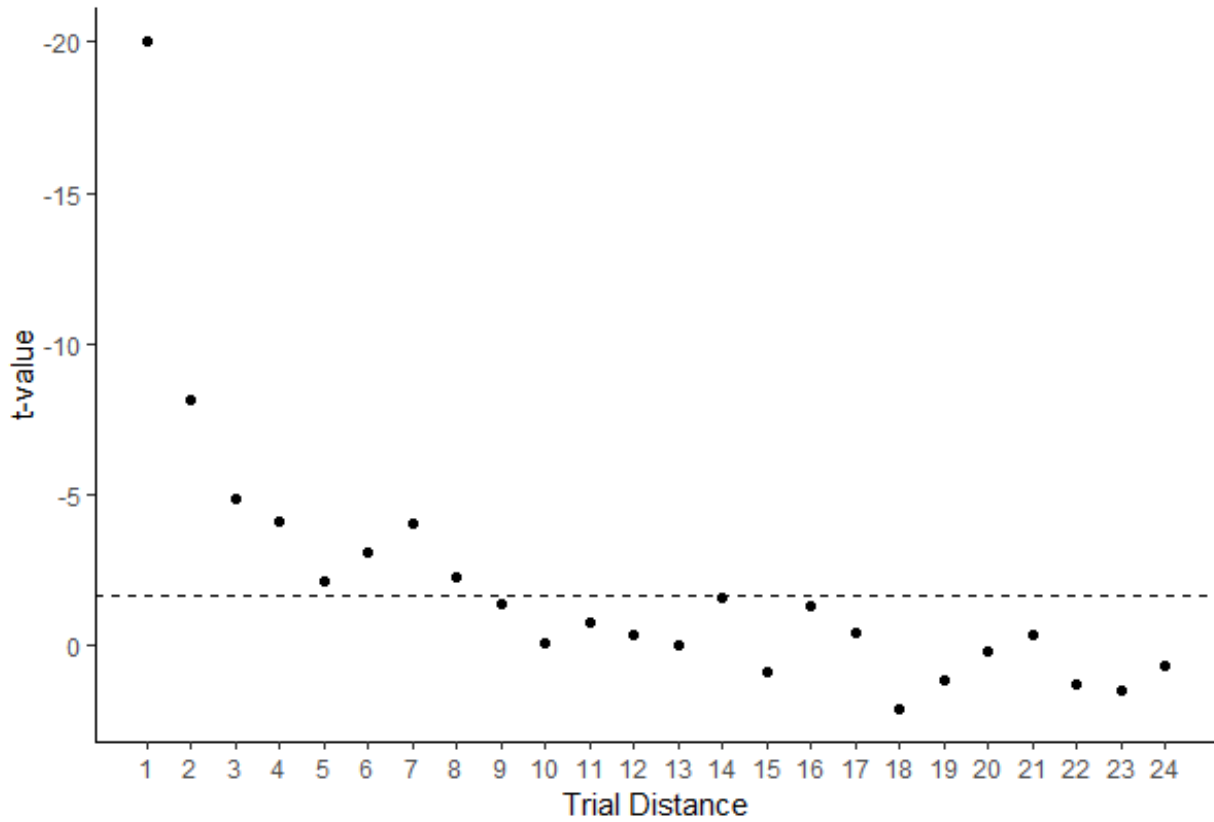
axis.

Figure 6. C_i C_{i-k} interaction t values of the model (*k* ranges from 1 - 24), plotted as a function of

trial distance. 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
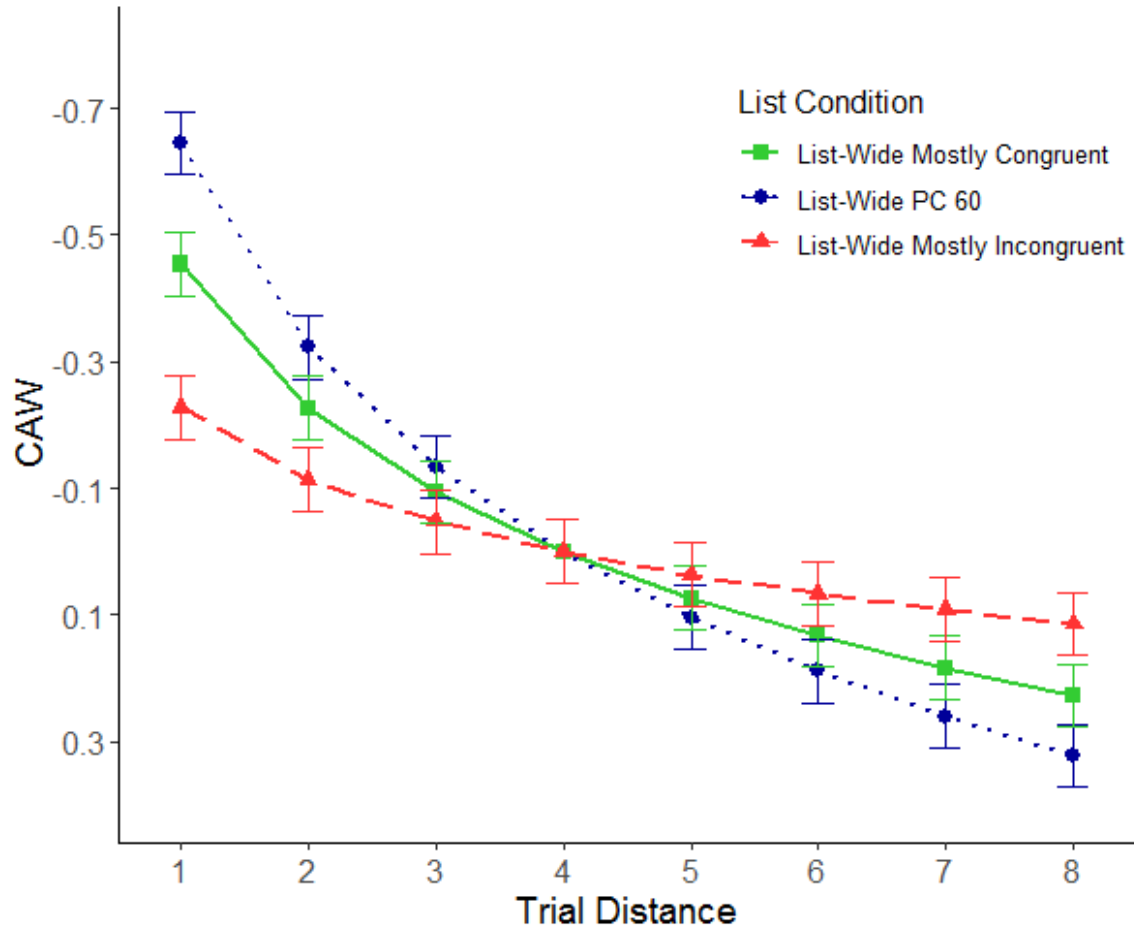
Gourley et al. (2016).

Figure 7. Estimates of the CAWs in Model 4 (full model) using data from Gourley et al. (2016).

Each condition is plotted after subtracting its intercept. The original scale of trial distance is
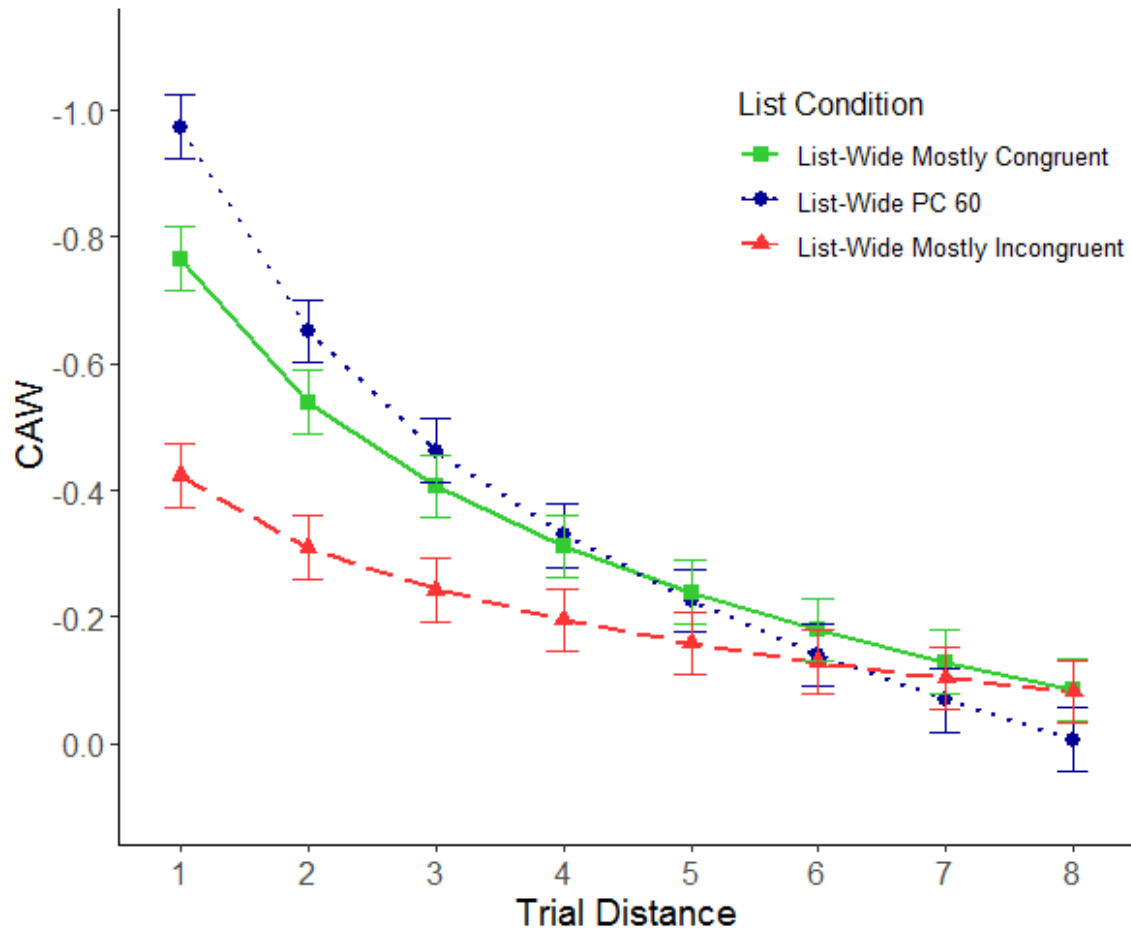
displayed on the x-axis.

Figure 8. Estimates of the CAWs in Model 4 (full model) using data from Gourley et al. (2016).

No intercept correction was applied. The original scale of trial distance is displayed on the x-
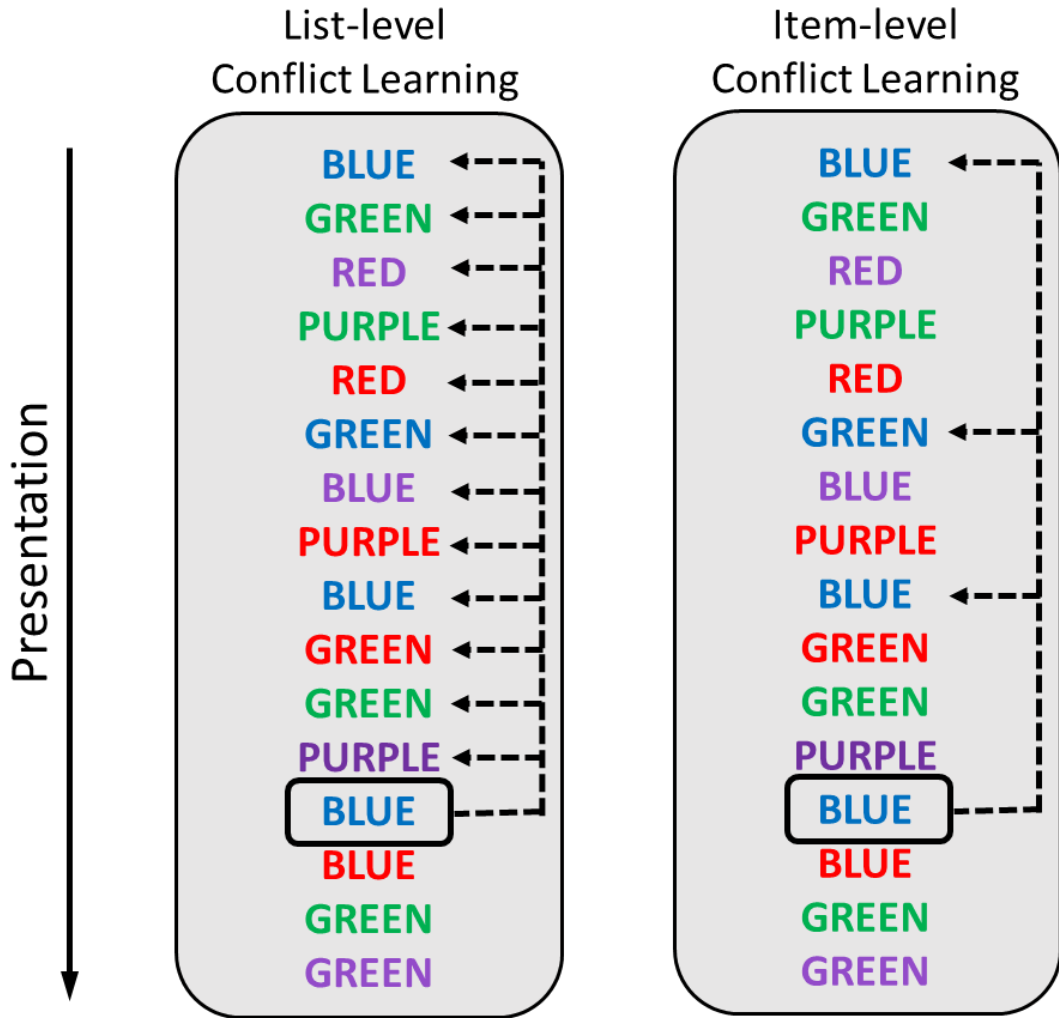
axis.

Figure 9. 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 (using an exponential weighting scheme) the conflict statuses of previous trials to arrive at 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 of previous instances of that item must be aggregated.

Figure 10. Estimates of the CAWs in Model 4 (full model) using mostly congruent and mostly incongruent items within the list-wide PC 50 condition in Gonthier et al. (2016). Each condition is plotted after subtracting its intercept. The original scale of trial distance is displayed on the x-axis.
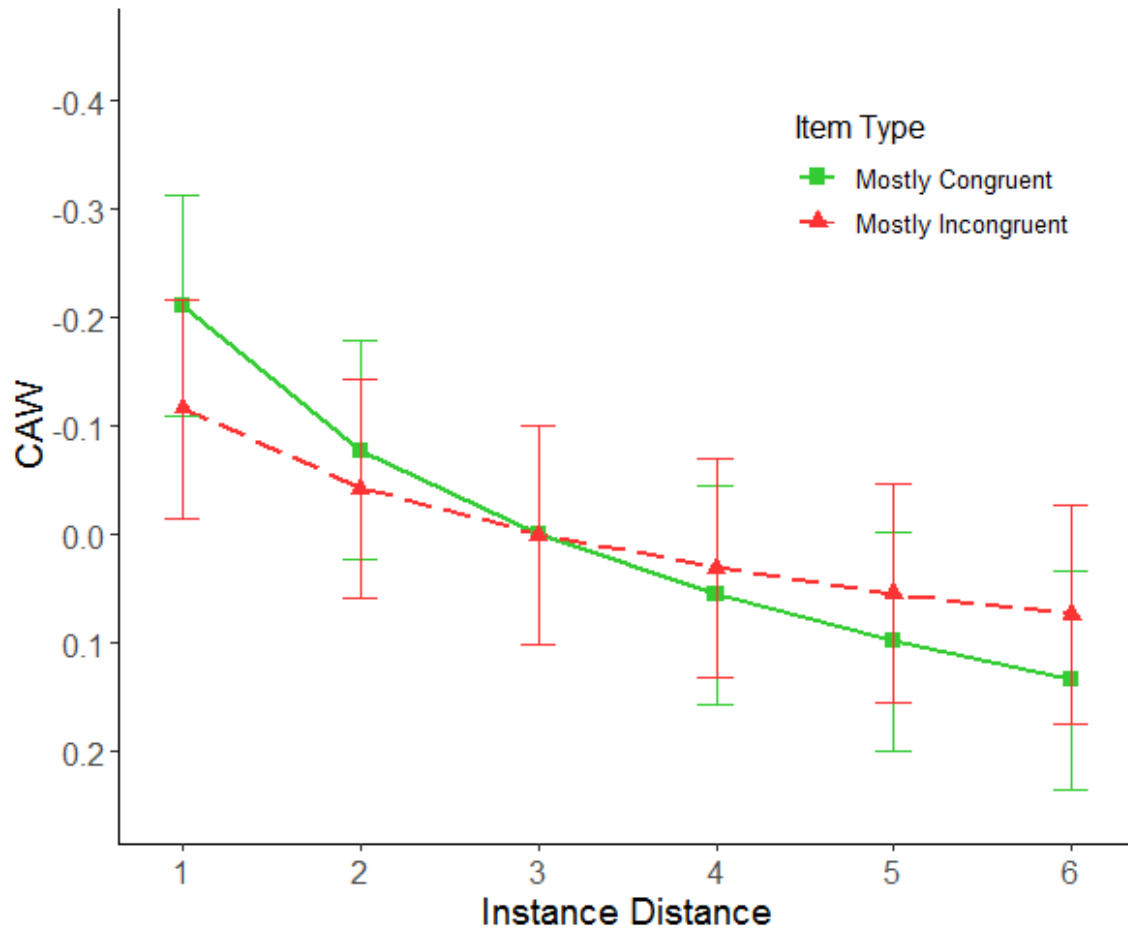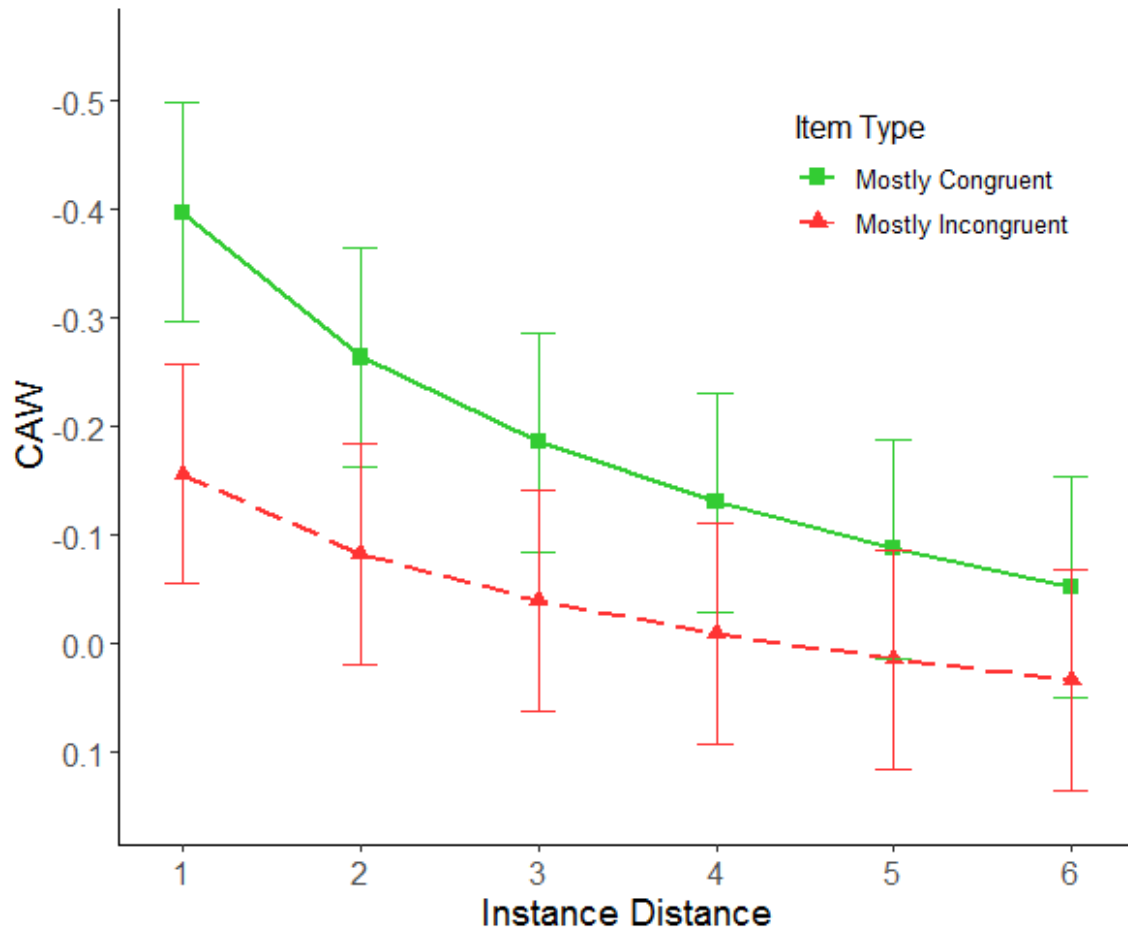
Figure 11. Estimates of the CAWs in Model 4 (full model) using mostly congruent and mostly incongruent items within the list-wide PC 50 condition in Gonthier et al. (2016). No intercept correction was applied. The original scale of trial distance is displayed on the x-axis.

Figure 12. Estimates of the CAWs in Model 4 (full model) using mostly congruent and mostly incongruent items within PC 50 lists from Bugg & Dey (2018). Each condition is plotted after subtracting its intercept. The original scale of trial distance is displayed on the x-axis.
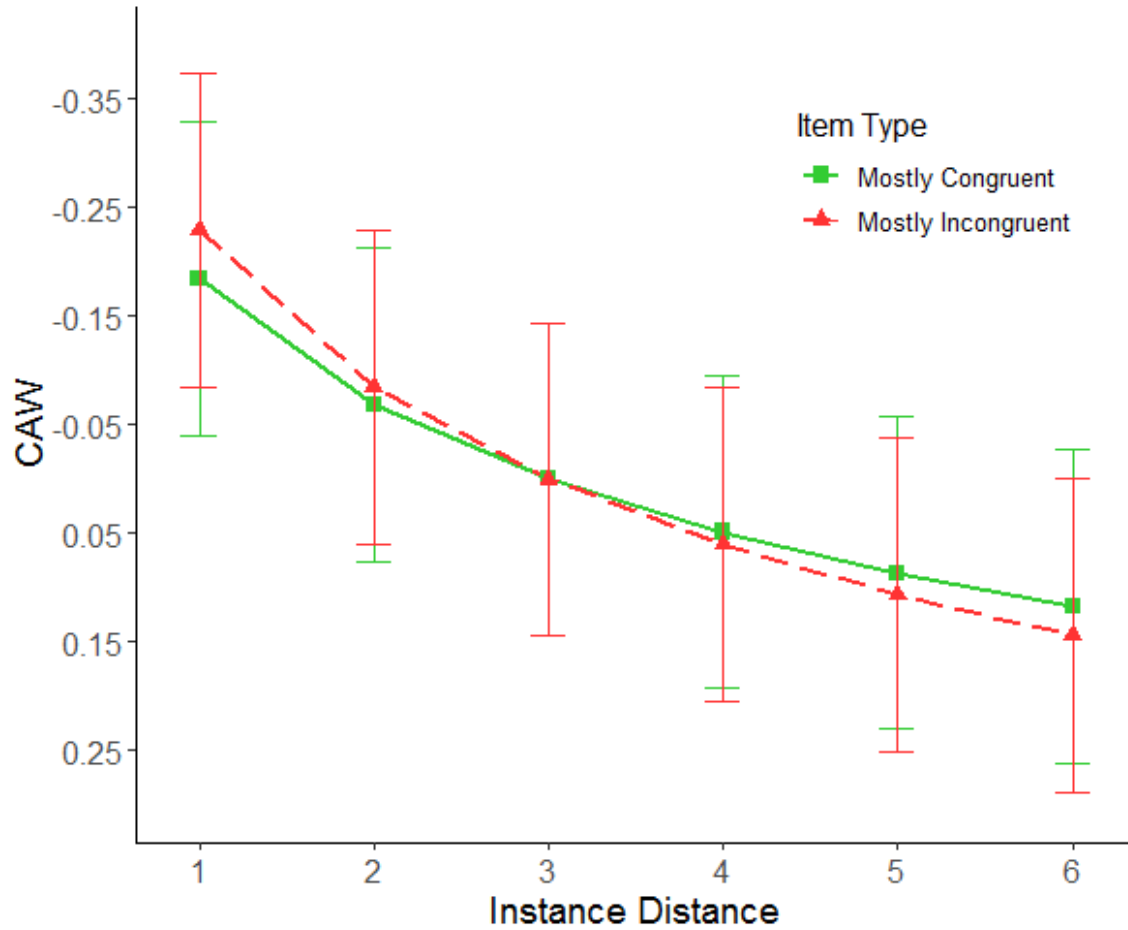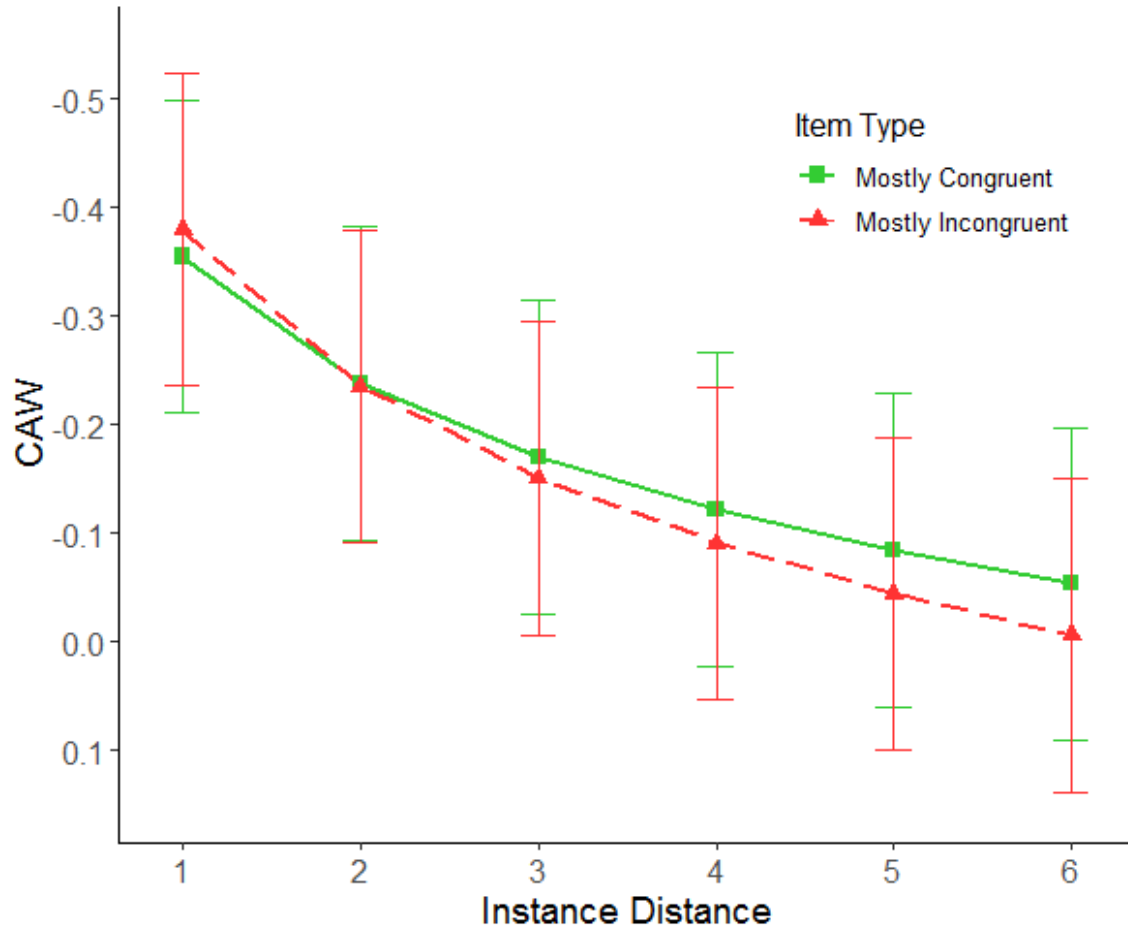
Figure 13. Estimates of the CAWs in Model 4 (full model) using mostly congruent and mostly incongruent items within PC 50 lists from Bugg & Dey (2018). No intercept correction was applied. The original scale of trial distance is displayed on the x-axis.

Table 1

Gonthier et al. (2017) List-Wide Proportion Congruence Model Comparisons

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	<.001
2. Condition	5	8058	-4024	2 vs. 0	22.46	<.001
3. Log trial distance + Condition	6	7961	-3974	3 vs. 1	23.04	<.001
				3 vs. 2	99.77	<.001
4. Log trial distance + Condition + Log trial distance x Condition	8	7960	-3971	4 vs. 3	5.40	0.067

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

Table 2

Gonthier et al. (2017) List-Wide Proportion Congruence Full Model Coefficients

Variable	B (SE)	t
(Intercept)	-.09 (.02)	-4.74
Log trial distance	.16 (.02)	6.62
LWMC condition	-.09 (.03)	-3.67
LWMI condition	.01 (.03)	.26
Log trial distance x LWMC condition	.01 (.03)	.19
Log trial distance x LWMI condition	-.06 (.03)	-1.91

Note. LWMC = list-wide mostly congruent; LWMI = list-wide mostly incongruent.

Table 3.

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

Model	df	AIC	log lik.	Test	X^2	p
0. Intercept	3	6103	-3049			
1. Log trial distance	4	5978	-2985	1 vs. 0	127.48	<.001
2. Condition	5	6095	-3042	2 vs. 0	12.44	.002
3. Log trial distance + Condition	6	5969	-2978	3 vs. 1	13.18	.001
				3 vs. 2	128.22	<.001
4. Log trial distance + Condition + Log trial distance x Condition	8	5953	-2969	4 vs. 3	19.74	<.001

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

Table 4.

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

Variable	*B* (*SE*)	*t*
(Intercept)	-.32 (.03)	-9.89
Log trial distance	.46 (.05)	9.72
LWMC condition	.02 (.04)	.382
LWMI condition	.13 (.04)	2.99
Log trial distance x LWMC condition	-.14 (.07)	-2.04
Log trial distance x LWMI condition	-.30 (.07)	-4.45

Note. LWMC = list-wide mostly congruent; LWMI = list-wide mostly incongruent.

Table 5.

Gonthier et al. (2016) Item-Specific Proportion Congruence Model Comparison

Model	*df*	AIC	log lik.	Test	X^2	*p*
0. Intercept	3	3529	-1762			
1. Log instance distance	4	3525	-1759	1 vs. 0	5.63	.018
2. Item type	4	3527	-1760	2 vs. 0	3.69	.054
3. Log trial distance + Condition	5	3523	-1757	3 vs. 1	3.71	.054
				3 vs. 2	5.64	.017
4. Log trial distance + Condition + Log trial distance x Condition	6	3525	-1757	4 vs. 3	.48	.488

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

Table 6.

Gonthier et al. (2016) Item-specific Proportion Congruence Full Model Coefficients

Variable	*B* (*SE*)	*t*
(Intercept)	-.19 (.06)	-3.22
Log instance distance	.19 (.09)	2.17
MI item	.15 (.08)	1.93
Log instance distance x MI item	-.08 (.13)	-.69

Note. MI = mostly incongruent. Mostly congruent items were used as the reference group.

Table 7.

Bugg & Dey (2018) Item-specific Proportion Congruence Model Comparisons

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	<.001
2. Item type	4	7322	-3657	2 vs. 0	.24	.618
3. Log trial distance + Condition	5	7288	-3639	3 vs. 1	.25	.615
				3 vs. 2	35.94	<.001
4. Log trial distance + Condition + Log trial distance x Condition	6	7290	-3639	4 vs. 3	.41	.522

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

Table 8.

Bugg & Dey (2018) Item-Specific Full Model Coefficients

Variable	*B* (*SE*)	*t*
(Intercept)	-.17 (.03)	-5.50
Log instance distance	.17 (.04)	3.80
MI item	.02 (.04)	.51
Log instance distance x MI item	.04 (.06)	.64

Note. MI = mostly incongruent. Mostly congruent items were used as the reference group.