

## Review

## Measuring Adaptive Control in Conflict Tasks

Senne Braem,<sup>1,\*</sup> Julie M. Bugg,<sup>2</sup> James R. Schmidt,<sup>3</sup> Matthew J.C. Crump,<sup>4</sup> Daniel H. Weissman,<sup>5</sup> Wim Notebaert,<sup>6</sup> and Tobias Egner<sup>7</sup>

The past two decades have witnessed an explosion of interest in the cognitive and neural mechanisms of adaptive control processes that operate in selective attention tasks. This has spawned not only a large empirical literature and several theories but also the recurring identification of potential confounds and corresponding adjustments in task design to create confound-minimized metrics of adaptive control. The resulting complexity of this literature can be difficult to navigate for new researchers entering the field, leading to suboptimal study designs. To remediate this problem, we present here a consensus view among opposing theorists that specifies how researchers can measure four hallmark indices of adaptive control (the congruency sequence effect, and list-wide, context-specific, and item-specific proportion congruency effects) while minimizing easy-to-overlook confounds.

### The Quest for Pure Measures of Adaptive Control

**Cognitive control** (see [Glossary](#)) allows people to act in ways that are consistent with their internal goals [1]. To investigate such control, psychologists often use selective attention tasks that create **conflict** by pitting instructed task goals against incompatible stimulus information and automatic action tendencies (i.e., conflict tasks). For example, in the seminal Stroop task [2], researchers study how the ability to identify the color of a printed word varies with whether the word cues a different semantic representation and response than the color (e.g., the word BLUE in red ink; incongruent trials) or whether these are the same (e.g., the word RED in red ink; congruent trials). Participants typically respond more slowly and less accurately on incongruent versus congruent trials. Researchers commonly consider the size of this ‘congruency effect’ as being indicative of the signal strength of the irrelevant dimension relative to the relevant dimension, as well as of the level of cognitive control applied – when congruency effects are relatively small, researchers infer that there is greater recruitment of cognitive control.

Importantly, conflict tasks also allow psychologists to study modulations of congruency effects that are thought to reflect adjustments of cognitive control; we refer to this as **adaptive control** (sometimes also called ‘control learning’ [3]). These dynamic adjustments of control are particularly important to measure because it is the matching of processing modes (e.g., a narrow vs a wide focus of attention) to changing environmental demands, and/or in response to performance monitoring signals (e.g., conflict), that characterizes adaptive behavior [4]. In other words, instead of conceptualizing control as a static, time-invariant process (e.g., by assessing mean congruency effects over an entire experiment), adaptive control research is concerned with how control is regulated in a dynamic and time-varying manner. This captures both the need to deal with a changing environment as well as the notion that control is costly and should be imposed only as much as necessary [5]. Research on adaptive control has already led to many important insights and influential theories [5–9], and continues to inspire an increasing number of studies. Moreover, beyond the basic research domain, adaptive control has been the topic of many studies and theories on developmental changes [10–12] and on various clinical disorders [13–19].

### Highlights

Early putative indices of adaptive control in conflict tasks have spurred not only a great deal of research but also numerous discussions on what these indices actually measure.

Recent studies have shown that adaptive control effects can be observed after controlling for low-level confounds. However, many canonical findings in the literature, for instance concerning the functional neuroanatomy of adaptive control, are based on older, confounded designs, and may thus be subject to revision.

This research field is now starting to experience a second wave of studies on adaptive control in conflict tasks employing improved designs that allow us to (re)address old and new questions.

<sup>1</sup>Vrije Universiteit Brussel, Brussels, Belgium

<sup>2</sup>Washington University in St. Louis, St. Louis, MO, USA

<sup>3</sup>Université Bourgogne Franche-Comté (UBFC), Dijon, France

<sup>4</sup>Brooklyn College of the City University of New York (CUNY), Brooklyn, NY, USA

<sup>5</sup>University of Michigan, Ann Arbor, MI, USA

<sup>6</sup>Ghent University, Ghent, Belgium

<sup>7</sup>Duke University, Durham, NC, USA

\*Correspondence: [senne.braem@vub.be](mailto:senne.braem@vub.be) (S. Braem).



However, numerous metrics of adaptive control have been put forth, criticized, and revised several times. Therefore, it can be difficult for new (or applied) researchers in this domain to infer what represents current best practices for studying adaptive control. In fact, many studies continue to use task designs or analysis strategies that researchers in the basic research community no longer consider to effectively measure adaptive control [20,21]. For example, a recent review on adaptive control in schizophrenia concluded that 'there are very few clearly interpretable studies on behavioral adaptation to conflict in the literature on schizophrenia' ([13], p. 209). Crucially, systematic comparisons between older, confound-prone, and newer, confound-minimized measures of adaptive control have shown differential behavioral effects [22] and patterns of brain activity [23, 24]. Therefore, the aim of the present paper is to promote best practices for investigating adaptive control, based on a current consensus view shared by different researchers in this field.

### The Need for 'Inducer' and 'Diagnostic' Items when Studying Adaptive Control

When employing conflict tasks, the goal of the researcher is typically to isolate changes in behavior that reflect adjustments to relatively abstract attentional settings or task representations (e.g., 'pay more attention to the target' or 'be cautious in selecting the response') as opposed to concrete settings (e.g., 'look at the green square' or 'press the left response key'). For example, in the Stroop task, abstract adjustments of control could involve paying more attention to the task-relevant color dimension or trying to inhibit the response cued by the task-irrelevant word dimension. Such adjustments are abstract in the sense that they should lead to generalizable performance benefits that are independent of specific stimulus features or actions [25]. For instance, increased attention to the task-relevant color dimension should lead to reduced interference from the task-irrelevant word dimension, regardless of the exact color and word that appears.

However, several researchers have pointed out that classic purported indices of adaptive control in conflict tasks can often be re-explained in terms of more basic stimulus–stimulus or stimulus–response learning processes ([21] for review). These considerations have led to various theoretical discussions about how such forms of lower-level learning relate to cognitive control (Box 1). However, experts in this domain generally agree that manipulations that promote learning at this concrete level are relatively easy to avoid. Therefore, if researchers want to study adaptive control independently of low-level learning, our recommendation is that they should employ paradigms that are designed to minimize opportunities for exploiting stimulus–response or stimulus–stimulus associations. We will refer to these design features in the remainder of this paper as 'confounds' (but see the section When Low-Level Learning Is Not a 'Confound').

To explain how to accomplish this goal, we brought together different researchers in this field (with different theoretical backgrounds, Box 1) to summarize an emerging consensus view on how to design conflict tasks to study adaptive control. It quickly became apparent that the best way to summarize our view is to emphasize one key experimental design principle that enables researchers to investigate adaptive control in a confound-minimized fashion. Specifically, the principle is to distinguish between **inducer items** that trigger adaptive control and **diagnostic items** that measure the effects of adaptive control on performance.

Without going into the intricacies of the different confounds when investigating adaptive control, which have been extensively discussed in other papers [21,26,27], we describe here how to create tasks that avoid these confounds. We focus on four common markers of adaptive control (Box 2 and Figure 1): the congruency sequence effect (CSE), the list-wide proportion congruency effect (LWPCE), the context-specific proportion congruency effect (CSPCE), and the item-specific proportion congruency effect (ISPCE). Specifically, we discuss how the inclusion of inducer and diagnostic items minimizes confounds that often prevent researchers from

### Glossary

**Adaptive control:** refers here to control processes or executive functions that dynamically adjust processing selectivity in response to changes in the environment or to internal (performance) monitoring signals (e.g., conflicts).

**Cognitive control:** the term (also 'executive functions') is generally used to describe a set of (not always well-defined) higher-order processes that are thought to direct, correct, and redirect behavior in line with internal goals and current context.

**(Cognitive) Conflict:** conflict in information processing is thought to occur when two or more mutually incompatible stimulus representations and/or response tendencies are triggered by a stimulus, such as an incongruent stimulus in the Stroop task (invoking e.g., both 'blue' and 'red').

**Conflict adaptation:** adaptive processes that are putatively triggered following the detection of conflict and are recruited for the purpose of resolving this conflict or preventing subsequent occurrences of conflict. This term is sometimes also used to refer specifically to the CSE.

**Contingency learning:** the general learning process of forming stimulus–stimulus and/or stimulus–response associations based on their co-occurrence, where the strength of the association increases as a function of the frequency of co-occurrence. While contingency learning is often discussed within the context of implicit learning, it is not necessary to make any assumptions about whether this learning occurs explicitly or implicitly, or is strategic or automatic.

**Diagnostic items:** items that are used to 'measure' the effects of adaptive control on performance. These items are sometimes also referred to as non-manipulated items, unbiased items, transfer items, or test items.

**Feature integration:** the idea that multiple features of a given stimulus are integrated or bound together in perception and memory. An extension of this idea holds that this integration of the features of an experience (or event) in memory also incorporates one's response to the stimulus into an episodic 'event file'. The subsequent presentation of one of those event features is then thought to facilitate the retrieval of the entire event file from memory. More recently, it has been proposed that event

Box 1. Cognitive Control versus Low-Level Learning

The observation that low-level learning contributes to various metrics of adaptive control has led to several theoretical discussions about how adaptive control works, and about how ‘low-level’ stimulus–stimulus or stimulus–response learning relates to cognitive control more broadly. Initially, the seminal conflict-monitoring account (Figure 1A) [6] proposed that conflict is detected by a conflict monitor which signals the need for adaptation to specific top-down task or control modules. However, some have subsequently argued that such modules are not necessary to explain adaptive control (Figure 1C) [21], whereas others have suggested that these top-down modules are recruited as a last-resort mechanism [50]. Further theories emphasize a close integration between low-level learning and cognitive control (Figure 1B,D) [25,100–102] to better account for observations such as context- and item-specific proportion congruency effects, or the domain-specificity of the congruency sequence effect ([44,103] for review). These theories suggest that control representations are embedded in the same associative network of stimulus and response representations. Some see an important role for conflict as a teaching signal that promotes these different forms of learning (Figure 1B) [101,102], whereas others have been agnostic about this [25], or have argued against conflict as a teaching signal in driving adaptive control (Figure 1D) [22].

Taken together, we believe these theories differ in their responses to two pertinent questions. First, some accounts consider that cognitive control and low-level learning rely on separate mechanisms or modules (Figure 1A,C), whereas others question this idea and capitalize on the similarities between the two by arguing for the associative learning or episodic binding of control states (Figure 1B,D). A second open question has been whether conflict serves as an active control signal (Figure 1A,B) or plays no necessary functional role (Figure 1C,D) in adaptive control.

Finally, it is important to mention two further issues. First, there are also various more-specific theories on the nature of the proposed conflict monitor [5,104–106] and the precise mechanisms that make up these control processes ([28,107], overviewed in [108]), that are beyond the scope of this article. Second, although the different measures discussed in this article are all thought to measure adaptive control beyond stimulus–response learning, they can still differ in their ‘degree’ of higher-order processing. For example, one interesting proposal is that adaptive control processes can also come about simply by learning one abstract property of the task, namely the time it takes to respond to the previous trial (i.e., temporal learning accounts [109,110]; but see [57,111]).

files may even include abstract attentional states, which could make it a form of adaptive control (Box 1).  
**Inducer items:** researchers manipulate these items to ‘trigger’ adaptive control. These items are sometimes also referred to as manipulated items, biased items, context items, or training items.

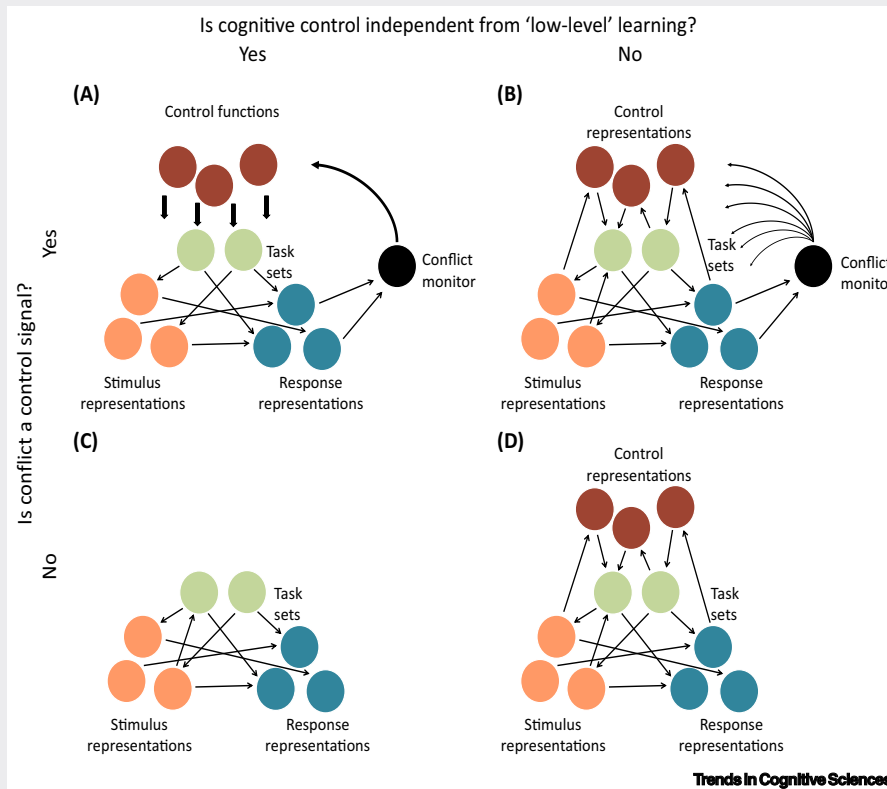


Figure 1. A Schematic Depiction of Four Major Groups of Theories of Adaptive Control.

**Box 2. Different Measures of Adaptive Control – What Do They Test?**

The most widely used marker of adaptive control in conflict tasks is the congruency sequence effect (CSE). The CSE refers to the observation that the congruency effect is smaller following incongruent versus congruent trials (see Figures 1A and 2 in main text) ([28], reviewed in [27,29]). The CSE is typically thought to measure a short-lived, reactive adaptation to a just-experienced conflict between competing response representations [6].

A second marker is the list-wide proportion congruency effect (LWPCE). The LWPCE refers to the observation that the congruency effect is smaller in blocks that contain mostly incongruent trials than in blocks that contain mostly congruent trials (Figures 1B and 2) ([28,112], reviewed in [26]). In contrast to the CSE, the LWPCE is usually considered to measure more global, and possibly proactive (i.e., anticipatory [113,114]), adaptations of control that accrue over a larger time-window than the previous trial. However, some researchers have suggested that these two effects might be mediated by the same underlying learning processes [6,35,102,115,116]. In fact, one can even consider that the typical LWPCE design is confounded by different proportions of previous-trial congruency, according to which the LWPCE could also reflect an accumulation of CSE effects. Although some studies have accounted for this, and still found a small LWPCE effect [117], the exact relation between these two types of adaptive control remains a topic of discussion.

For anyone new to the field who is interested in assessing adaptive control, we recommend focusing on the CSE or LWPCE because these effects have been studied most extensively with the here-described confound-minimized designs.

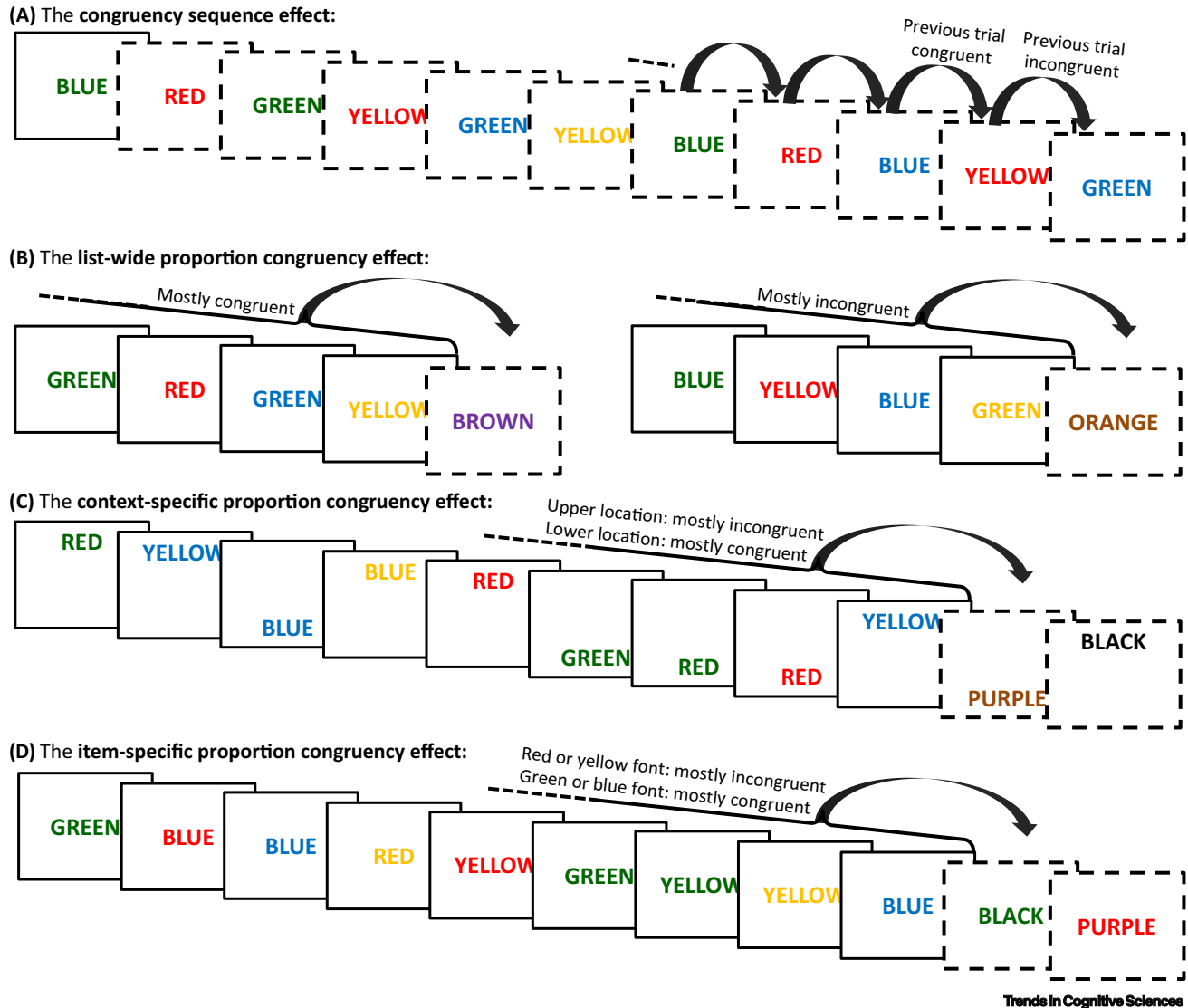
The third and fourth ways to manipulate proportion congruency are to link it to specific contexts or specific items, respectively. In tasks measuring the context-specific proportion congruency effect (CSPCE) [58], the proportions of congruent and incongruent trials are tied to a contextual feature that can change on a trial-by-trial basis and that is not part of the task stimulus itself (e.g., the spatial location of Stroop stimuli; Figure 1C). In tasks measuring the item-specific proportion congruency effect (ISPCE) [71], the proportions of congruent and incongruent trials are contingent on a specific task-relevant feature (e.g., color in the Stroop task; Figure 1D). Similar to the LWPCE, both the CSPCE and the ISPCE are thought to reflect adaptive control processes that learn about proportion congruency across several trials, rather than only the previous trial as for the CSE. In contrast to the LWPCE, and more like the CSE, however, these adaptive control processes are recruited reactively after stimulus onset in the CSPCE and ISPCE.

Importantly, please note that all original observations of these four effects came from designs that are now considered confounded because low-level learning (e.g., contingency learning, associative learning, episodic memory of stimulus–response episodes) could explain the modulation of the congruency effect.

accurately assessing adaptive control using these four measures. We use the Stroop task as an example throughout this paper. However, our recommendations also apply to other conflict tasks (Box 3).

**The Congruency Sequence Effect**

The CSE, sometimes referred to as the ‘Gratton effect’ [28] or ‘**conflict adaptation** effect’ [6], describes the finding that the congruency effect is reduced following incongruent versus congruent trials ([27,29] for review). The CSE is thought to measure adaptive control on a trial-by-trial basis (Box 2). However, over a decade ago researchers noted that, in typical two-alternative forced choice (2-AFC) conflict tasks with small stimulus sets, the nature of stimulus or response repetitions across consecutive trials differs for the different conditions that researchers use to calculate the CSE. Thus, they argued that **feature integration** processes, stemming from such unequal repetitions, rather than adaptive control processes, might engender the CSE [30, 31]. Later, researchers also noted that, in most studies using 4-AFC tasks, which are helpful for avoiding stimulus and response repetitions, each possible distractor is paired more often with the congruent target than with each possible incongruent target, which can lead to **contingency learning** [32]. Such tasks confound different congruency conditions with different contingency conditions and, hence, different congruency sequences with different contingency sequences. Thus, it is not possible to interpret the CSE as an index of adaptive control in such tasks (although note that some do consider contingency sequence learning a form of adaptive control [33,34], albeit not one that involves conflict processing; also Box 1). Reacting to such findings, researchers began to use separate inducer items to trigger adaptive control, and diagnostic items to measure the effects of adaptive control [30,35–37].



**Figure 1. Different Indices of Adaptive Control in the Stroop Task.** For each measure, broken screens indicate the use of diagnostic items to study adaptive changes in behavior that were triggered by different items. Note that for the examples (B–D) diagnostic items are usually randomly presented among the other items (not only at the end of a run), which is here not the case for reasons of space. (A) The congruency sequence effect measures adaptive changes following incongruent versus congruent trials. (B) The list-wide proportion congruency effect measures adaptations in control in blocks where mostly congruent trials are presented (left) versus blocks where mostly incongruent trials are presented (right). (C) The context-specific proportion congruency effect investigates adaptive control in contexts associated with mostly incongruent items (e.g., stimuli presented at the top of the screen) versus contexts with mostly congruent items (e.g., stimuli presented at the bottom of the screen). (D) The item-specific proportion congruency effect probes adaptations in control to item features that are presented under mostly incongruent conditions (i.e., the ink colors red and yellow) versus item features presented in mostly congruent conditions (i.e., the ink colors blue and green).

One increasingly popular way to employ inducer and diagnostic items is to create two stimulus sets (and associated response sets), each with their own congruent and incongruent items, and to alternate between these sets on a trial-by-trial basis [35–38]. For instance, in the Stroop task, one might present (i) blue and green color and word stimuli in odd trials, and (ii) red and yellow color and word stimuli in even trials (Figure 1A). Crucially, by alternating between these two stimulus sets, the design guarantees that the inducer items that trigger adaptive control in one trial are different from the diagnostic items that measure adaptive control in the next trial, thereby ensuring

**Box 3. Different Conflict Tasks – Which One Should I Choose?**

We use the Stroop task [2] as an example throughout our paper because it is the most popular of all conflict tasks. However, researchers have also used other paradigms to study adaptive control. For instance, in the flanker task participants identify a central target while ignoring congruent or incongruent flanking distractors. As another example, in the Simon task participants make a lateralized response (e.g., a left button press) to identify a non-spatial feature of a stimulus (e.g., its color – red) while ignoring the location of the stimulus on the screen, which can be congruent (e.g., left) or incongruent (e.g., right).

The same signatures of adaptive control (such as CSE, LWPCE, etc.) have been found across different types of conflict tasks. However, there is currently little evidence that adaptive control can cross between conflict tasks for the CSE (i.e., where the inducer trial is from one task, and the diagnostic from another; reviewed in [44,103]), and only some for the LWPCE [56,118,119]. Further, indices of adaptive control within different conflict tasks show low to no correlations with one another [85]. Thus, one should think carefully about which conflict task to employ. In so doing, at least two issues should be considered.

First, researchers posit that different tasks invoke different types of conflict (and adaptive control processes) as a function of the cognitive processing stage(s) in which there is an overlap between task-relevant and task-irrelevant stimulus features [120]. For instance, whereas conflict in the Stroop task likely involves clashing semantic features, conflict in the Simon task likely involves location-triggered response priming. One may also consider employing tasks that invoke other types of conflict, such as emotional conflict [121].

Second, different conflict tasks may have different degrees of power to detect changes in congruency effects. For example, although the CSE can be reliably observed in the Stroop task ([35,85], but see [38]), its effect size is relatively small (e.g.,  $N = 178$ ,  $\eta^2 = 0.05$  [85]). By contrast, tasks wherein the distractor precedes the target, such as the prime-probe task, produce a CSE that has a larger effect size [37]. Similarly, Stroop or flanker tasks wherein the task-irrelevant information appears before target presentation also result in larger CSE effects [122]. Therefore, we strongly recommend that statistical power analysis is carried out before setting up an experiment and that the sample size and the number of trials are chosen accordingly (main text).

that the CSE must reflect higher-level cognitive adjustments that are independent of stimulus–response learning. Instead of employing this alternating-sets design, one can also consider using a large set of stimuli and responses such that each trial is a new stimulus (e.g., a picture-naming task with a large set of picture–word interference stimuli; [39] experiment 1b).

Three additional aspects deserve attention when studying the CSE. First, to avoid contingency learning confounds, one should present each unique congruent stimulus and each unique incongruent stimulus in each stimulus set equally often (e.g., the word RED in yellow should occur as often as the word RED in red [32,40]). Thus, for example, congruent trials should occur 50% of the time when alternating between a pair of 2-AFC tasks, 33% of the time when alternating between a pair of 3-AFC tasks, and so on.

Second, while the two stimulus sets should contain different stimuli and responses, the task should be the same for both (e.g., color categorization) because task switching reduces the CSE. That is, when using different ‘stimulus sets’, the idea should be to create non-overlapping sets of stimuli and responses that are all part of the same task set (e.g., name ink colors, ignore words), unless, of course, one wishes to investigate the influence of task switching on the CSE ([41–43]; reviewed in [44]). Using different response sets may reduce the CSE in some situations [36,45,46], especially when using arbitrary stimulus–response mappings [47]. However, in some tasks, such as the prime-probe task, it is possible to observe robust CSEs with both arbitrary [38] and non-arbitrary [37] stimulus–response mappings provided that the participants perceive that all the stimuli and responses belong to the same task set.

Third, one should always exclude all trials that follow an unusual event. This is because the CSE is crucially dependent on what occurred on the previous trial. For instance, we typically exclude all trials following an error (in addition to current-trial errors) and the first trial of each block.

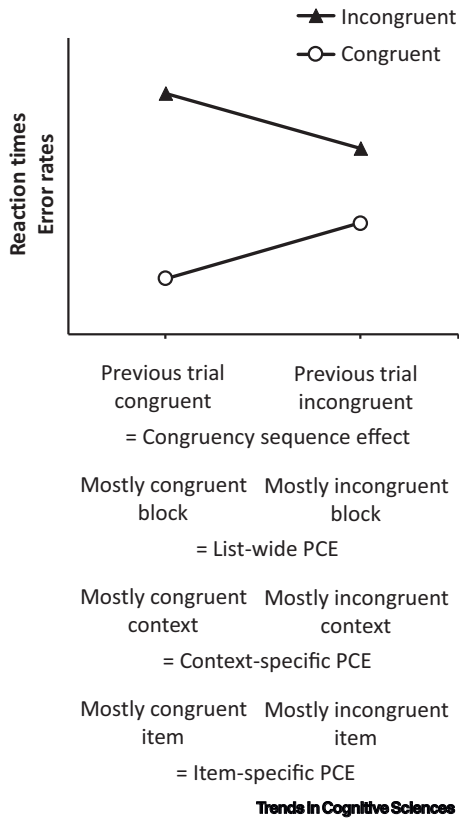


Figure 2. Behavioral Indices of Adaptive Control. The four performance indices are characterized by a reduction in the congruency effect either following incongruent versus congruent trials (i.e., congruency sequence effect), or in mostly incongruent versus mostly congruent conditions (i.e., proportion congruency effects, PCEs). The interaction depicted in this figure represents a generic form – the actual form of the interaction may vary from one index to another.

Crucially, in our consensus view, following the procedures described in this section will allow researchers to better attribute the CSE (Figure 2) to an adaptive control process, rather than to feature integration or contingency learning mechanisms. Note that there are still various alternative accounts about what exactly this adaptive control process entails (Box 1), but they all involve mechanisms that go beyond low-level stimulus–response learning.

### The List-Wide Proportion Congruency Effect

The LWPCE describes the finding of a smaller congruency effect in blocks of more- relative to less-frequent incongruent trials. The LWPCE is thought to measure global adaptations of control by the likelihood of experiencing conflict (high or low) in a particular block (list) of trials ([48] for a user’s guide to the proportion congruency manipulation). In many earlier studies, however, the frequency of incongruent trials was confounded with the frequency with which specific stimuli appeared. For example, a smaller congruency effect in mostly incongruent (MI) blocks than in mostly congruent (MC) blocks could reflect either adaptive control, or merely having seen specific incongruent stimuli (i.e., BLUE in red) more frequently in the MI block, because participants typically respond more quickly to stimuli that appear more frequently as a result of low-level learning.

The best approach to studying the LWPCE while minimizing confounds also involves creating two sets of stimuli (and associated responses). The inducer items manipulate the relative frequencies of congruent and incongruent trials. By contrast, the diagnostic items measure the effect of the inducer items on the LWPCE. The diagnostic items typically consist of equal percentages of

congruent and incongruent trials, although the percentages could be different, provided that they remain the same in the MI and MC blocks (e.g., [49]). The inducer and diagnostic items are then randomly intermixed. As a practical example, during MI blocks the inducer items might consist of 80% incongruent trials and 20% congruent trials, whereas the diagnostic items might consist of 50% incongruent and 50% congruent trials (Figure 1B). During MC blocks, the proportion congruency would be reversed for the inducer items (e.g., 20% incongruent, 80% congruent) but remain the same for the diagnostic items (i.e., 50% congruent).

We recommend two additional procedures when investigating the LWPCE. First, to maximize the list-wide bias engendered by the inducer items in each block (i.e., such that overall PC is low in the MI block and high in the MC block), inducer items should appear more frequently than diagnostic items ([50] for sample frequencies).

Second, the inducer set should ideally be composed of at least three different stimuli and associated responses (e.g., many studies have used a Stroop task with four different words and their corresponding colors, resulting in three equally frequent response options on incongruent inducer trials [50–52]). This is because manipulating proportion congruency on small sets of stimuli (e.g., two colors and their corresponding words in a Stroop task) promotes contingency learning, which may weaken the triggering of adaptive control based on the proportion congruency of the list ([53,54]; but see [55,56]). This second recommendation naturally increases the number of responses in the task. For example, four colors for inducer items and two colors for diagnostic items would require six response options. These are readily accommodated in vocal versions of the Stroop task. Alternatively, one can pretrain participants on a six alternative forced-choice (6-AFC) manual (i.e., button-pressing) conflict task.

Using such designs, we can make inferences about list-wide (global) adaptive control by analyzing the inducer and diagnostic items separately. If the congruency effect is reduced for the diagnostic items in MI blocks compared to MC blocks, this suggests that changes in control have taken place that cannot be attributed to the frequency with which individual items appeared (Boxes 1 and 2 for different accounts). Finally, we note that it is also possible to investigate the LWPCE using a design wherein a unique item appears on each trial (e.g., new pictures in a picture–word Stroop), which bypasses the need for two different stimulus sets [57].

### The Context-Specific Proportion Congruency Effect

The CSPCE refers to the observation that the congruency effect can change when proportion congruency is manipulated across two or more contexts that vary on a trial-by-trial basis. For instance, presenting a higher proportion of incongruent stimuli in one of two possible stimulus locations can lead to smaller congruency effects at that location than at the other location. Unlike the LWPCE, the CSPCE is thought to index adaptations to different congruency proportions within a block which are predicted by task-irrelevant contextual features that are not part of the main task (e.g., stimulus location [58], color [59], shape surrounding the stimulus [60], temporal presentation windows [61], etc.). Relatively few studies have determined whether the CSPCE appears when measured with diagnostic items, and the evidence is mixed ([62–65]; but see [66]). Consequently, there is an ongoing controversy about whether the CSPCE reflects adaptive control or a form of contingency learning [67]. Regardless, we believe that, for researchers who want to use the CSPCE as a measure of adaptive control, the same design rules apply as for the CSE and the LWPCE.

Specifically, researchers should employ distinct sets of inducer and diagnostic items. Importantly, both sets should appear in both contexts (within the same block), but the inducer



items should vary in proportion congruency (and thereby trigger adaptive control) depending on the context. For example, in [Figure 1C](#) the relevant color blue appears most often with an incongruent word distractor in the upper (MI) screen location but with a congruent word distractor in the lower (MC) screen location. The diagnostic items (e.g., 50% congruent in each location) measure the effects of adaptive control. Crucially, if the congruency effect for the diagnostic items is smaller in the MI context (e.g., in the upper screen location) than in the MC context (e.g., in the lower screen location, [Figure 2](#)), one may conclude that there is evidence for adaptive control.

### The Item-Specific Proportion Congruency Effect

Our final index of adaptive control is the ISPCE, which refers to the finding that the size of the congruency effect for a particular item varies with how frequently it appears in incongruent versus congruent trials. For example, the congruency effect is smaller for target items that appear more frequently with incongruent distractors than for target items that appear more frequently with congruent distractors. Much like the LWPCE and the CSPCE, the ISPCE reflects adaptations to different proportions of congruent and incongruent trials. However, it differs from these other indices of adaptive control in that researchers manipulate the proportion congruency of different task-relevant items (e.g., target colors [\[68\]](#) or pictures [\[69,70\]](#)). The original studies aimed to manipulate the proportion congruency of the distractors [\[71\]](#). However, it is preferable that the relevant stimulus feature is predictive of proportion congruency because, when the irrelevant feature is predictive (e.g., the word in the Stroop task), the ISPCE that results can be driven by contingency learning, as shown in several studies ([\[69,72\]](#); but see [\[69\]](#) experiment 1, and [\[68\]](#) experiment 3, four-item set condition). Accordingly, when examining the ISPCE in the Stroop task, researchers manipulate proportion congruency across the different print colors that appear randomly in each block (e.g., blue is MC, red is MI). Crucially, unlike for the other three indices of adaptive control, the ‘diagnostic items’ should share the predictive feature (e.g., color) with the biased inducer items. Otherwise, item-specific adjustments cannot be assessed.

Design-wise, examining the ISPCE begins with the creation of two stimulus sets that vary in their task-relevant feature (e.g., color in the Stroop task). Unlike for the LWPCE and CSPCE, these stimulus sets both consist of inducer items; that is, the triggers of adaptive control. In the Stroop task, for example, one set of colors (e.g., blue and green) appears more frequently with a congruent word (MC items), whereas the other set of colors (e.g., red and yellow) appears more frequently with an incongruent word (MI items). During the task these items are randomly intermixed ([Figure 1D](#)). Further, the item sets overlap, meaning, in our example, that the colors red and yellow would appear not only with the words ‘red’ and ‘yellow’ but also with the words ‘blue’ and ‘green’, and vice versa for the colors blue and green. This might discourage participants from relying on contingency learning based on the distractor word to predict the target response [\[69\]](#), which can occur when the sets do not overlap ([\[71,72\]](#), but see [\[68\]](#) experiment 3, four-item set condition). One typically tests for the presence of an ISPCE by determining whether the congruency effect is smaller for MI (e.g., red and yellow) items than for MC (e.g., green and blue) items. However, although the ISPCE in this design is based on differences in proportion congruency associated with the relevant dimension, the distractor words are not completely uncorrelated with responses ([\[23,69\]](#) experiment 2, and [\[68\]](#) experiments 1 and 2, [\[70,73\]](#)).

Therefore, consistent with our general recommendation, we encourage researchers to create a third set of diagnostic items to assess the ISPCE without item-frequency differences. Two types of approaches currently use diagnostic items, but each comes with a cautionary note.

First, one can create diagnostic items by choosing novel items from the same categories as the inducer items ([69] experiment 2, also [70]). For example, in a Stroop task wherein participants name a picture category (e.g., dogs, birds) while ignoring a superimposed word (e.g., cat), one can use a novel set of exemplars as diagnostic items (e.g., new dogs and birds that are 50% congruent). Importantly, while this approach already reduces some prominent contingency learning confounds, it could still be influenced by smaller remaining distractor-response contingencies or contingency learning at a categorical level based on the distractor word ([74,75], reviewed in [21]). No study has established that these remaining contingency learning opportunities can give rise to the ISPCE in these particular designs, but researchers should be aware of this caveat.

A second approach is to use diagnostic items that involve new distractor features (e.g., words). Crucially, this approach controls for all currently known frequency and contingency learning confounds. For example, in the Stroop task, diagnostic trials consist of the MC and MI inducer colors paired equally often with incongruent non-inducer distractor words such as 'purple' or 'black' (Figure 1D). Using this design, one can examine whether responses are faster for (i) MI inducer colors paired with non-inducer incongruent words, than for (ii) MC inducer colors paired with non-inducer incongruent words. Such a result indicates adaptive control in the absence of frequency and contingency learning confounds. However, analogously to the CSPCE, only one study has investigated item-specific control using this particular approach ([68] experiment 2). Therefore, while we can confidently recommend this approach, it will be important to assess its robustness in future studies.

### Power Analysis and Design Planning

Our recommendation to use inducer and diagnostic items for measuring adaptive control also comes with a warning regarding statistical power. The CSE, LWPCE, ISPCE, and CSPCE are all calculated as the difference between two congruency effects, which are themselves difference scores. Taking a difference between difference scores can increase variability [76]. Hence, if a design has sufficient power to measure a congruency effect, it might not necessarily have sufficient power to measure a difference between congruency effects. Moreover, splitting a design into inducer and diagnostic items can reduce the number of trials used to measure the effect of interest (with the exception of the CSE), thus further reducing statistical power.

Therefore, we recommend that researchers ensure their design has sufficient trials and participants to achieve a desired level of statistical power for assessing adaptive control. To aid with power calculations (and determining the required sample size) for adaptive control designs, Crump and Brosowsky [77] created `conflictPower` (<https://crumplab.github.io/conflictPower/>), a free R package for estimating statistical power using Monte Carlo simulations (using the method from [64]). This approach provides statistical power estimates for congruency effects and differences between congruency effects that depend on the number of participants, effect size, the number of trials per cell, and estimates about the shape of underlying reaction-time distributions using ex-Gaussian parameters.

We are sensitive to the fact that acquiring more trials or bigger sample sizes can sometimes be challenging, especially in clinical studies. Therefore, we are not recommending that patient studies be withheld for the sole reason that they might lack the statistical power to draw firm conclusions. One way to address this general issue (that is not specific to this literature; e.g., [78]) is to rely more on meta-analyses or to conduct multilaboratory studies for accurate inferences. Another could be the development of new paradigms that account for the same (or new) confounds, but that focus more on detecting reliable individual differences.

### When Low-Level Learning Is Not a 'Confound'

We recommend that researchers employ diagnostic items as a means to measure adaptive control because effects on these items are not easily explained by simple response-repetition effects or the formation of stimulus–response (or stimulus–stimulus) associations. Therefore, they license consideration of adaptive control accounts. This also led us to label as 'confounds' design features that allow these lower-level associative effects to influence performance, consistent with the decades-old literature on this topic. Importantly, however, this terminology is not intended to imply that stimulus–response or stimulus–stimulus learning is not adaptive in its own right – it very certainly is. In fact, as noted in [Box 1](#), some recent theories have begun to emphasize that adaptive control might rely on the same learning processes as lower-level forms of learning, the only difference being that they typically operate on more abstract representations (e.g., task difficulty, congruency expectancy, error likelihood). Similarly, we do not want to argue that adaptive control cannot act at the level of specific stimulus or response features. It most likely can, and such specificity could even be considered to be a marker of its adaptivity rather than a confound. However, demonstrating adaptive control at this level (where it is probably most powerful) comes with the problem that we cannot distinguish it from other low-level explanations (for a partial exception, see the ISPCE section).

Of course, in most real-world situations, it is likely that adaptations of cognitive control parameters (such as task focus) go hand in hand with lower-level learning about the specifics of the environment. This notion is inherent in the study of the CSPCE and ISPCE discussed above, where concrete context or stimulus features condition control processes. Moreover, this has also been investigated in studies on the interaction between lower-level learning and the CSE [22,79–82]. Accordingly, we would like to emphasize that, although confound-minimized designs isolate adaptive control processes from lower-level learning processes, we do not view adaptive control as being either context-free or untethered from lower-level processing. On the contrary, we believe that context and specific stimulus features may be potent drivers of, and interact with, control.

Finally, our recommendations do not imply that there is no value in employing the type of task designs we denote as 'confounded' in this article – what constitutes a confound clearly depends on the intention and hypotheses of the researcher. For example, a researcher might have a hypothesis that is agnostic about the mechanism that underlies the effect, have diverging predictions depending on the underlying process (e.g., adaptive control versus contingency learning), or simply be interested in the interaction between these different levels of learning [22,79–82]. Instead, our argument is that, if it is the goal of the researcher to isolate the effects of adaptive control from concurrent effects of more concrete and lower-level learning processes (as is often the case), then the use of the above-described inducer-diagnostic item design is strongly recommended.

### Concluding Remarks

The current literature on adaptive control is characterized by a wide heterogeneity of paradigms and designs. Thus, for researchers who are not 'in the weeds' of this field, it can be difficult to infer a consensus view on the steps that are necessary to optimally study adaptive control in conflict tasks. We argue here that there is one key consideration. Specifically, to measure the hypothetical effects of adaptive control, researchers should employ diagnostic stimuli (and associated responses) that do not overlap with inducer stimuli (and associated responses) that trigger such control (for a partial exception, see the ISPCE section).

Of course, creating optimal conditions for measuring adaptive control is not limited to the use of inducer and diagnostic items. For example, researchers should also ensure that they have

### Outstanding Questions

Will previous findings hold when using the recommended measures of adaptive control? For example, topics such as the domain-generalty of control processes, their reward-sensitivity, the role of consciousness, and the relation to other constructs such as working memory capacity have mostly been studied using designs that potentially measured stimulus–response or stimulus–stimulus learning.

Are the learning processes that mediate adaptive control qualitatively different from, or similar to, the learning processes that mediate stimulus–response or stimulus–stimulus learning?

What learning processes support transfer and generalization of control within and across contexts and items?

Which brain regions or processes are involved in adaptive control when using inducer-diagnostic item designs, and how do they relate to the neural mechanisms that underlie stimulus–response or stimulus–stimulus learning?

What role, if any, does the detection of conflict play in driving adaptive control? For example, is there a functional role for conflict in triggering adaptive control processes?

Can future designs be further modified to detect more reliable individual differences within and across tasks? – and what are the correlates of individual differences in adaptive control?

sufficient statistical power, which will depend on the exact conflict task and effect under investigation. It is also worth noting that indices of adaptive control often show a low correlation across, and even within, different conflict tasks (with or without confounds [83–85]), similar to other measures of cognitive control [86,87]. In other words, while most of these effects can be reliably observed at the group level, they are not optimized for studying interindividual differences. This could also explain why markers of adaptive control often show low or null correlations with other related measures such as working memory capacity (e.g., [88]). In part, this can be expected for tasks that are popular precisely because they produce reliable group effects, and thus low between-subject variance [86,89].

More broadly, although this paper has focused on four prominent measures of adaptive control in conflict tasks (Box 4 for other measures in conflict tasks), we believe that the same ideas apply to the study of adaptive control in other tasks. For example, consider studies of task switching. In this case the repetition of task cues [90,91] and stimulus–response features [92] from one trial to the next influences the switch cost, which is thought to reflect, at least partly, the time needed by control processes to reconfigure task sets [93]. Interestingly, similar measures of adaptive control exist in the task-switching literature, such as the list-wide [94], context-specific [95], and item-specific [96,97] modulations of the switch cost (reviewed in [98]). However, while some task-switching studies have circumvented such low-level repetition confounds by presenting different cues or stimuli on each trial [99], very few have employed confound-minimized designs [92].

In conclusion, the past two decades have witnessed an exponentially increasing interest in the study of adaptive control. During this period, specific indices of adaptive control in conflict tasks and the methods for avoiding experimental confounds have been extensively discussed and revised numerous times. For those outside the field it can be difficult to follow the current optimal means to study these control processes. In this paper we have presented a consensus view that emphasizes the inclusion of inducer and diagnostic items in the experimental design. We believe that future studies using such ‘confound-minimized’ designs (in isolation or in combination with other factors, such as contingency learning) will allow researchers to address and revisit exciting research questions (see Outstanding Questions) and substantially advance our understanding of adaptive control.

#### Box 4. More Measures of Adaptive Control in Conflict Tasks

This paper focuses on four popular indices of adaptive control in conflict tasks, but there are also other interesting measures, especially when using variations of conflict tasks such as conflict task-switching studies or free-choice paradigms. First, although most indices of adaptive control come from single-task paradigms, some can also be observed in (conflict) task-switching paradigms (reviewed in [123]). A seminal finding in this domain is that the switch cost (i.e., the performance cost associated with alternating versus repeating the just-performed task) is greater when the previous trial was incongruent versus congruent [124]. This effect is thought to measure enhanced processing of task-relevant information after an incongruent trial as a result of adaptive control processes (e.g., [125,126]; but see [92]), similar to some explanations of the CSE [6].

A second group of studies have begun to look at adaptations to conflict in free-choice tasks. All measures discussed thus far come from tasks wherein participants are clearly instructed to perform one task or another. However, in an attempt to create more ecologically valid paradigms, some researchers have started to employ voluntary task-switching paradigms wherein participants choose which task to perform in each trial [127]. Using such paradigms, some researchers have studied whether people are more likely to repeat a task, or switch to a new task, if the previous trial was incongruent or congruent [128,129]. Similarly, others have studied whether people, when given a choice, tend to avoid or approach contexts that are associated with a higher proportion of incongruent trials [60,130]. Many of these studies have demonstrated that people tend to prefer congruent over incongruent trials (and associated tasks), which is consistent with the idea that adaptive control is also costly and demanding [5].

### Acknowledgments

We would like to thank Luis Jiménez, Juan Lupiáñez, and three anonymous reviewers for useful suggestions on an earlier version of this manuscript, as well as Frederick Verbruggen for inspiration for this project. J.M.B. was supported by the National Institute On Aging of the National Institutes of Health under award number R21AG057937, J.R.S. was supported by the Programme d'Investissements d'Avenir (project ISITE-BFC) under award number ANR15-IDEX-0003, and T.E. was supported in part by the National Institutes of Health under award number 2R01 MH087610. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health.

### References

- Miller, E.K. and Cohen, J.D. (2001) An integrative theory of prefrontal cortex function. *Annu. Rev. Neurosci.* 24, 167–202
- Stroop, J.R. (1935) Studies of interference in serial verbal reactions. *J. Exp. Psychol.* 18, 643
- Chiu, Y.C. and Egner, T. (2019) Cortical and subcortical contributions to context-control learning. *Neurosci. Biobehav. Rev.* 99, 33–41
- Goschke, T. (2003) Voluntary action and cognitive control from a cognitive neuroscience perspective. In *Voluntary Action: Brains, Minds, and Sociality* (Maasen, S. et al., eds), pp. 49–85, Oxford University Press
- Shenhav, A. et al. (2013) The expected value of control: an integrative theory of anterior cingulate cortex function. *Neuron* 79, 217–240
- Botvinick, M.M. et al. (2001) Conflict monitoring and cognitive control. *Psychol. Rev.* 108, 624–652
- Braver, T.S. (2012) The variable nature of cognitive control: a dual mechanisms framework. *Trends Cogn. Sci.* 16, 106–113
- Kerns, J.G. et al. (2004) Anterior cingulate conflict monitoring and adjustments in control. *Science* 303, 1023–1026
- Ridderinkhof, K.R. et al. (2004) The role of the medial frontal cortex in cognitive control. *Science* 306, 443–447
- Bugg, J.M. (2014) Evidence for the sparing of reactive cognitive control with age. *Psychol. Aging* 29, 115–127
- Iani, C. et al. (2014) Response inhibition and adaptations to response conflict in 6- to 8-year-old children: evidence from the Simon effect. *Atten. Percept. Psychophys.* 76, 1234–1241
- Larson, M.J. et al. (2012) Cognitive control and conflict adaptation similarities in children and adults. *Dev. Neuropsychol.* 37, 343–357
- Abrahamse, E. et al. (2016) Conflict adaptation in schizophrenia: reviewing past and previewing future efforts. *Cogn. Neuropsychiatry* 21, 197–212
- Clawson, A. et al. (2013) Cognitive control adjustments and conflict adaptation in major depressive disorder. *Psychophysiology* 50, 711–721
- Lansbergen, M.M. et al. (2007) Stroop interference and attention-deficit/hyperactivity disorder: a review and meta-analysis. *Neuropsychology* 21, 251
- Larson, M.J. et al. (2012) Cognitive control and conflict adaptation in youth with high-functioning autism. *J. Child Psychol. Psychiatry* 53, 440–448
- Stuedte-Schmiedgen, S. et al. (2014) Trauma exposure is associated with increased context-dependent adjustments of cognitive control in patients with posttraumatic stress disorder and healthy controls. *Cogn. Affect. Behav. Neurosci.* 14, 1310–1319
- Tulek, B. et al. (2013) Attentional control is partially impaired in obstructive sleep apnea syndrome. *J. Sleep Res.* 22, 422–429
- Wylie, S.A. et al. (2010) The effect of Parkinson's disease on the dynamics of on-line and proactive cognitive control during action selection. *J. Cogn. Neurosci.* 22, 2058–2073
- Schmidt, J.R. (2013) Questioning conflict adaptation: proportion congruent and Gratton effects reconsidered. *Psychon. Bull. Rev.* 20, 615–630
- Schmidt, J.R. (2019) Evidence against conflict monitoring and adaptation: an updated review. *Psychon. Bull. Rev.* 26, 753–771
- Weissman, D.H. et al. (2016) Different levels of learning interact to shape the congruency sequence effect. *J. Exp. Psychol. Learn. Mem. Cogn.* 42, 566
- Chiu, Y.C. et al. (2017) The caudate nucleus mediates learning of stimulus-control state associations. *J. Neurosci.* 37, 1028–1038
- Jiang, J. et al. (2015) Memory meets control in hippocampal and striatal binding of stimuli, responses, and attentional control states. *J. Neurosci.* 35, 14885–14895
- Egner, T. (2014) Creatures of habit (and control): a multi-level learning perspective on the modulation of congruency effects. *Front. Psychol.* 5, 1247
- Bugg, J.M. and Crump, M.J. (2012) In support of a distinction between voluntary and stimulus-driven control: a review of the literature on proportion congruent effects. *Front. Psychol.* 3, 367
- Duthoo, W. et al. (2014) The heterogeneous world of congruency sequence effects: an update. *Front. Psychol.* 5, 1001
- Gratton, G. et al. (1992) Optimizing the use of information: strategic control of activation of responses. *J. Exp. Psychol. Gen.* 121, 480
- Egner, T. (2007) Congruency sequence effects and cognitive control. *Cogn. Affect. Behav. Neurosci.* 7, 380–390
- Mayr, U. et al. (2003) Conflict adaptation effects in the absence of executive control. *Nat. Neurosci.* 6, 450
- Hommel, B. et al. (2004) A feature-integration account of sequential effects in the Simon task. *Psychol. Res.* 68, 1–17
- Schmidt, J.R. and De Houwer, J. (2011) Now you see it, now you don't: controlling for contingencies and stimulus repetitions eliminates the Gratton effect. *Acta Psychol.* 138, 176–186
- Jiménez, L. et al. (2009) Sequential congruency effects in implicit sequence learning. *Conscious. Cogn.* 18, 690–700
- Schmidt, J.R. et al. (2007) Contingency learning without awareness: evidence for implicit control. *Conscious. Cogn.* 16, 421–435
- Jiménez, L. and Méndez, A. (2014) Even with time, conflict adaptation is not made of expectancies. *Front. Psychol.* 5, 1042
- Kim, S. and Cho, Y.S. (2014) Congruency sequence effect without feature integration and contingency learning. *Acta Psychol.* 149, 60–68
- Schmidt, J.R. and Weissman, D.H. (2014) Congruency sequence effects without feature integration or contingency learning confounds. *PLoS One* 9, e102337
- Weissman, D.H. et al. (2014) Determinants of congruency sequence effects without learning and memory confounds. *J. Exp. Psychol. Hum. Percept. Perform.* 40, 2022
- Duthoo, W. et al. (2014) The congruency sequence effect 3.0: a critical test of conflict adaptation. *PLoS one* 9, e110462
- Mordkoff, J.T. (2012) Observation: three reasons to avoid having half of the trials be congruent in a four-alternative forced-choice experiment on sequential modulation. *Psychon. Bull. Rev.* 19, 750–757
- Hazeltine, E. et al. (2011) The boundaries of sequential modulations: evidence for set-level control. *J. Exp. Psychol. Hum. Percept. Perform.* 37, 1898
- Kiesel, A. et al. (2006) Evidence for task-specific resolution of response conflict. *Psychon. Bull. Rev.* 13, 800–806
- Notebaert, W. and Verguts, T. (2008) Cognitive control acts locally. *Cognition* 106, 1071–1080
- Braem, S. et al. (2014) What determines the specificity of conflict adaptation? A review, critical analysis, and proposed synthesis. *Front. Psychol.* 5, 1134
- Braem, S. et al. (2011) Conflict adaptation by means of associative learning. *J. Exp. Psychol. Hum. Percept. Perform.* 37, 1662

46. Janczyk, M. and Leuthold, H. (2018) Effector system-specific sequential modulations of congruency effects. *Psychon. Bull. Rev.* 25, 1066–1072
47. Lim, C.E. and Cho, Y.S. (2018) Determining the scope of control underlying the congruency sequence effect: roles of stimulus–response mapping and response mode. *Acta Psychol.* 190, 267–276
48. Bugg, J.M. (2017) Context, conflict, and control. In *The Wiley Handbook of Cognitive Control* (Egner, T., ed.), pp. 79–96, John Wiley & Sons
49. Hutchison, K.A. (2011) The interactive effects of listwide control, item-based control, and working memory capacity on Stroop performance. *J. Exp. Psychol. Learn. Mem. Cogn.* 37, 851
50. Bugg, J.M. (2014) Conflict-triggered top-down control: default mode, last resort, or no such thing? *J. Exp. Psychol. Learn. Mem. Cogn.* 40, 567
51. Bugg, J.M. and Chanani, S. (2011) List-wide control is not entirely elusive: evidence from picture-word Stroop. *Psychon. Bull. Rev.* 18, 930–936
52. Gonthier, C. et al. (2016) Dissociating proactive and reactive control in the Stroop task. *Mem. Cogn.* 44, 778–788
53. Bugg, J.M. et al. (2008) Multiple levels of control in the Stroop task. *Mem. Cogn.* 36, 1484–1494
54. Blais, C. and Bunge, S. (2010) Behavioral and neural evidence for item-specific performance monitoring. *J. Cogn. Neurosci.* 22, 2758–2767
55. Schmidt, J.R. (2017) Time-out for conflict monitoring theory: preventing rhythmic biases eliminates the list-level proportion congruent effect. *Can. J. Exp. Psychol./Rev. Can. Psychol. Exp.* 71, 52
56. Wühr, P. et al. (2015) Generalizing attentional control across dimensions and tasks: evidence from transfer of proportion-congruent effects. *Q. J. Exp. Psychol.* 68, 779–801
57. Spinelli, G. et al. (2019) Adaptation to conflict frequency without contingency and temporal learning: evidence from the picture-word interference task. *J. Exp. Psychol. Hum. Percept. Perform.* Published online May 30, 2019. <http://doi.org/10.1037/xhp0000656>
58. Crump, M.J.C. et al. (2006) The context-specific proportion congruent Stroop effect: location as a contextual cue. *Psychon. Bull. Rev.* 13, 316–321
59. Lehle, C. and Hübner, R. (2008) On-the-fly adaptation of selectivity in the flanker task. *Psychon. Bull. Rev.* 15, 814–818
60. Schouppe, N. et al. (2014) Context-specific control and context selection in conflict tasks. *Acta Psychol.* 146, 63–66
61. Wendt, M. and Kiesel, A. (2011) Conflict adaptation in time: foreperiods as contextual cues for attentional adjustment. *Psychon. Bull. Rev.* 18, 910–916
62. Bejjani, C. et al. (2018) Control by association: transfer of implicitly primed attentional states across linked stimuli. *Psychon. Bull. Rev.* 25, 617–626
63. Crump, M.J.C. and Milliken, B. (2009) The flexibility of context-specific control: evidence for context-driven generalization of item-specific control settings. *Q. J. Exp. Psychol.* 62, 1523–1532
64. Crump, M.J.C. et al. (2017) Reproducing the location-based context-specific proportion congruent effect for frequency unbiased items: a reply to Hutcheon and Spieler (2016). *Q. J. Exp. Psychol.* 70, 1792–1807
65. Crump, M.J.C. et al. (2018) Context-dependent control of attention capture: evidence from proportion congruent effects. *Can. J. Exp. Psychol./Rev. Can. Psychol. Exp.* 72, 91–104
66. Hutcheon, T.G. and Spieler, D.H. (2017) Limits on the generalizability of context-driven control. *Q. J. Exp. Psychol.* 70, 1292–1304
67. Schmidt, J.R. and Lemerrier, C. (2019) Context-specific proportion congruent effects: compound-cue contingency learning in disguise. *Q. J. Exp. Psychol.* 72, 1119–1130
68. Bugg, J.M. and Hutchison, K.A. (2013) Converging evidence for control of color-word Stroop interference at the item level. *J. Exp. Psychol. Hum. Percept. Perform.* 39, 433
69. Bugg, J.M. et al. (2011) Why it is too early to lose control in accounts of item-specific proportion congruency effects. *J. Exp. Psychol. Hum. Percept. Perform.* 37, 844
70. Bugg, J.M. and Dey, A. (2018) When stimulus-driven control settings compete: on the dominance of categories as cues for control. *J. Exp. Psychol. Hum. Percept. Perform.* 44, 1905
71. Jacoby, L.L. et al. (2003) Item-specific control of automatic processes: Stroop process dissociations. *Psychon. Bull. Rev.* 10, 638–644
72. Schmidt, J.R. and Besner, D. (2008) The Stroop effect: why proportion congruent has nothing to do with congruency and everything to do with contingency. *J. Exp. Psychol. Learn. Mem. Cogn.* 34, 514
73. Bugg, J.M. (2012) Dissociating levels of cognitive control: The case of Stroop interference. *Curr. Dir. Psychol. Sci.* 21, 302–309
74. Schmidt, J.R. et al. (2018) Category learning in the color-word contingency learning paradigm. *Psychon. Bull. Rev.* 25, 658–666
75. Schmidt, J.R. and De Houwer, J. (2016) Contingency learning tracks with stimulus–response proportion. *Exp. Psychol.* 63, 79–88
76. Cronbach, L.J. and Furby, L. (1970) How we should measure 'change': or should we? *Psychol. Bull.* 74, 68
77. Crump, M.J.C. and Brosowsky, N.P. (2019) conflictPower: simulation based power analysis for adaptive control designs. R package version 0.1.0. <https://crumplab.github.io/conflictPower/>
78. Verbruggen, F. et al. (2019) A consensus guide to capturing the ability to inhibit actions and impulsive behaviors in the stop-signal task. *eLife* 8, e46323
79. Braem, S. et al. (2014) Reward determines the context-sensitivity of cognitive control. *J. Exp. Psychol. Hum. Percept. Perform.* 40, 1769
80. Dignath, D. et al. (2019) Reconciling cognitive-control and episodic-retrieval accounts of sequential conflict modulation: binding of control-states into event-files. *J. Exp. Psychol. Hum. Percept. Perform.* (in press)
81. Pires, L. et al. (2018) Cognitive control during a spatial Stroop task: comparing conflict monitoring and prediction of response-outcome theories. *Acta Psychol.* 189, 63–75
82. Spapé, M.M. and Hommel, B. (2008) He said, she said: episodic retrieval induces conflict adaptation in an auditory Stroop task. *Psychon. Bull. Rev.* 15, 1117–1121
83. Feldman, J.L. and Freitas, A.L. (2016) An investigation of the reliability and self-regulatory correlates of conflict adaptation. *Exp. Psychol.* 63, 237–247
84. Ruitenberg, M. et al. (2019) Learning to be in control involves response-specific mechanisms. *Atten. Percept. Psychophys.* Published online May 9, 2019. <https://doi.org/10.3758/s13414-019-01753-0>
85. Whitehead, P.S. et al. (2019) Are cognitive control processes reliable? *J. Exp. Psychol. Learn. Mem. Cogn.* 45, 765–778
86. Hedge, C. et al. (2018) The reliability paradox: why robust cognitive tasks do not produce reliable individual differences. *Behav. Res. Methods* 50, 1166–1186
87. Paap, K.R. and Sawi, O. (2016) The role of test–retest reliability in measuring individual and group differences in executive functioning. *J. Neurosci. Methods* 274, 81–93
88. Meier, M.E. and Kane, M.J. (2013) Working memory capacity and Stroop interference: global versus local indices of executive control. *J. Exp. Psychol. Learn. Mem. Cogn.* 39, 748
89. Enkavi, A.Z. et al. (2019) Large-scale analysis of test-retest reliabilities of self-regulation measures. *Proc. Natl. Acad. Sci.* 116, 5472–5477
90. Logan, G.D. and Bundesen, C. (2003) Clever homunculus: is there an endogenous act of control in the explicit task-cuing procedure? *J. Exp. Psychol. Hum. Percept. Perform.* 29, 575
91. Mayr, U. and Kliegl, R. (2000) Task-set switching and long-term memory retrieval. *J. Exp. Psychol. Learn. Mem. Cogn.* 26, 1124–1140
92. Schmidt, J.R. and Liefvooghe, B. (2016) Feature integration and task switching: diminished switch costs after controlling for stimulus, response, and cue repetitions. *PLoS one* 11, e0151188
93. Monsell, S. (2003) Task switching. *Trends Cogn. Sci.* 7, 134–140

94. Dreisbach, G. and Haider, H. (2006) Preparatory adjustment of cognitive control in the task switching paradigm. *Psychon. Bull. Rev.* 13, 334–338
95. Crump, M.J.C. and Logan, G.D. (2010) Contextual control over task-set retrieval. *Atten. Percept. Psychophys.* 72, 2047–2053
96. Chiu, Y.C. and Egner, T. (2017) Cueing cognitive flexibility: item-specific learning of switch readiness. *J. Exp. Psychol. Hum. Percept. Perform.* 43, 1950
97. Leboe, J.P. et al. (2008) Probe-specific proportion task repetition effects on switching costs. *Percept. Psychophys.* 70, 935–945
98. Braem, S. and Egner, T. (2018) Getting a grip on cognitive flexibility. *Curr. Dir. Psychol. Sci.* 27, 470–476
99. Schneider, D.W. (2015) Isolating a mediated route for response congruency effects in task switching. *J. Exp. Psychol. Learn. Mem. Cogn.* 41, 235
100. Abrahamse, E. et al. (2016) Grounding cognitive control in associative learning. *Psychol. Bull.* 142, 693
101. Blais, C. et al. (2007) Item-specific adaptation and the conflict-monitoring hypothesis: a computational model. *Psychol. Rev.* 114, 1076–1086
102. Verguts, T. and Notebaert, W. (2008) Hebbian learning of cognitive control: dealing with specific and nonspecific adaptation. *Psychol. Rev.* 115, 518
103. Egner, T. (2008) Multiple conflict-driven control mechanisms in the human brain. *Trends Cogn. Sci.* 12, 374–380
104. Botvinick, M.M. (2007) Conflict monitoring and decision making: reconciling two perspectives on anterior cingulate function. *Cogn. Affect. Behav. Neurosci.* 7, 356–366
105. Alexander, W.H. and Brown, J.W. (2011) Medial prefrontal cortex as an action-outcome predictor. *Nat. Neurosci.* 14, 1338
106. Silvetti, M. et al. (2011) Value and prediction error in medial frontal cortex: integrating the single-unit and systems levels of analysis. *Front. Hum. Neurosci.* 5, 75
107. Ridderinkhof, R.K. (2002) Micro- and macro-adjustments of task set: activation and suppression in conflict tasks. *Psychol. Res.* 66, 312–323
108. Weissman, D.H. et al. (2017) Identifying stimuli that cue multiple responses triggers the congruency sequence effect independent of response conflict. *J. Exp. Psychol. Hum. Percept. Perform.* 43, 677
109. Schmidt, J.R. (2013) Temporal learning and list-level proportion congruency: conflict adaptation or learning when to respond? *PLoS One* 8, e82320
110. Schmidt, J.R. and Weissman, D.H. (2016) Congruency sequence effects and previous response times: conflict adaptation or temporal learning? *Psychol. Res.* 80, 590–607
111. Cohen-Shikora, E.R. et al. (2018) Assessing the temporal learning account of the list-wide proportion congruence effect. *J. Exp. Psychol. Learn. Mem. Cogn.* Published online December 27, 2018. <https://doi.org/10.1037/xlm0000670>
112. Logan, G.D. and Zbrodoff, N.J. (1979) When it helps to be misled: facilitative effects of increasing the frequency of conflicting stimuli in a Stroop-like task. *Mem. Cogn.* 7, 166–174
113. Aben, B. et al. (2019) Context-dependent modulation of cognitive control involves different temporal profiles of fronto-parietal activity. *NeuroImage* 189, 755–762
114. De Pisapia, N. and Braver, T.S. (2006) A model of dual control mechanisms through anterior cingulate and prefrontal cortex interactions. *Neurocomputing* 69, 1322–1326
115. Aben, B. et al. (2017) Beyond trial-by-trial adaptation: a quantification of the time scale of cognitive control. *J. Exp. Psychol. Hum. Percept. Perform.* 43, 509
116. Jiang, J. et al. (2014) Bayesian modeling of flexible cognitive control. *Neurosci. Biobehav. Rev.* 46, 30–43
117. Torres-Quesada, M. et al. (2014) Gradual proportion congruent effects in the absence of sequential congruent effects. *Acta Psychol.* 149, 78–86
118. Funes, M.J. et al. (2010) Sustained vs. transient cognitive control: evidence of a behavioral dissociation. *Cognition* 114, 338–347
119. Torres-Quesada, M. et al. (2013) Dissociating proportion congruent and conflict adaptation effects in a Simon–Stroop procedure. *Acta Psychol.* 142, 203–210
120. Kornblum, S. et al. (1990) Dimensional overlap: cognitive basis for stimulus–response compatibility – a model and taxonomy. *Psychol. Rev.* 97, 253
121. Egner, T. et al. (2007) Dissociable neural systems resolve conflict from emotional versus nonemotional distracters. *Cereb. Cortex* 18, 1475–1484
122. Bombek, K. et al. (2017) The role of temporal predictability for early attentional adjustments after conflict. *PLoS one* 12, e0175694
123. Schuch, S. et al. (2019) Monitoring and control in multitasking. *Psychon. Bull. Rev.* 26, 222–240
124. Goschke, T. (2000) Intentional reconfiguration and J-TI involuntary persistence in task set switching. *Control of cognitive processes: Attention and performance XVIII*, 18 p. 331
125. Brown, J.W. et al. (2007) A computational model of fractionated conflict-control mechanisms in task-switching. *Cogn. Psychol.* 55, 37–85
126. Braem, S. et al. (2012) Reward modulates adaptations to conflict. *Cognition* 125, 324–332
127. Arrington, C.M. and Logan, G.D. (2004) The cost of a voluntary task switch. *Psychol. Sci.* 15, 610–615
128. Orr, J.M. et al. (2012) The influence of response conflict on voluntary task switching: a novel test of the conflict monitoring model. *Psychol. Res.* 76, 60–73
129. Dignath, D. et al. (2015) Flexible conflict management: conflict avoidance and conflict adjustment in reactive cognitive control. *J. Exp. Psychol. Learn. Mem. Cogn.* 41, 975
130. Desender, K. et al. (2017) Avoiding the conflict: metacognitive awareness drives the selection of low-demand contexts. *J. Exp. Psychol. Hum. Percept. Perform.* 43, 1397