The Boundaries of Adaptive Control

by

Lauren D. Grant

A dissertation submitted in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
(Psychology)
in the University of Michigan
2021

Doctoral Committee:

Professor Daniel H. Weissman, Chair
Assistant Professor David Brang
Assistant Professor Taraz Lee
Assistant Professor Stephan F. Taylor
Assistant Professor Michael Vesia

Lauren D. Grant

ldgran@umich.edu

ORCID iD: 0000-0002-2546-8473

Acknowledgements

I would like to thank all of my committee members for their feedback and support. Your thoughtful comments were critical for the success of my final dissertation studies. Special thanks to Daniel, my chair, for embarking on this exciting journey with me. Your support, advice, and enthusiasm have greatly shaped my development as a researcher.

I am grateful for my family, who are a constant source of love and support. I am also incredibly grateful for the support of my partner, Aaron. Your humor, kindness, and support has helped keep me sane this past year. Lastly, I am grateful for the support of my friends Emily, Christine, Ting Ting, and Pia, to name just a few.

I would also like to thank my undergraduate research assistants in the lab, without whom my studies would have never come to fruition. Special thanks to Samantha Cerpa, who helped with the writing of Chapter 5. Your inquisitiveness and willingness to learn constantly remind me of the joy of research and teaching!

Table of Contents

List of Tables

List of Figures

Abstract

Cognitive control processes that enable purposeful behavior are often context-specific. An individual, for example, may inhibit the tendency to speak loudly at a restaurant but not at a party. However, the nature of contextual boundaries for cognitive control processes remains unclear. For instance, it remains unclear as to why repeating contextual features enhances the congruency sequence effect (CSE), a common index of adaptive control, whereas alternating such features reduces this effect. To address such ambiguities, we sought to distinguish among various (and often conflicting) hypotheses concerning the boundaries of the CSE. In Chapters 2 and 3, we investigated the role of target-defining features and task sets in engendering or amplifying the CSE, respectively. Based on our findings, in Chapters 4 and 5 we explored whether task sets or episodic contexts determine the boundaries of the CSE. Interestingly, both studies suggested that task sets, rather than episodic context alone, determine the boundaries of the CSE. Thus, in Chapter 6, we further investigated the role of task sets in the context of episodic retrieval and action control. Contrary to prior studies, our findings suggested that non-hierarchical bindings between features – as opposed to hierarchical task sets – provide an explanation for task set boundaries of the CSE. Taken together, these studies provide a comprehensive overview on the boundaries of adaptive control.

Chapter 1 Introduction

In everyday life, individuals must flexibly adapt to an ever-changing environment. For instance, consider the situation of two individuals attempting to cross an intersection. Upon pressing the button for the pedestrian walk sign, they realize the sign is non-functional. Thus, to safely cross the intersection, they shift their attention to the traffic light. This ability to flexibly shift from one source of information (walk sign) to another (traffic light) is an integral aspect of adaptive control.

To investigate adaptive control processes in the laboratory, researchers often employ a variety of distractor-interference tasks such as the Stroop, Simon, and Flanker tasks (B. A. Eriksen & Eriksen, 1974; Stroop, 1935; Simon & Rudell, 1967; C. W. Eriksen & Schultz, 1979). For instance, consider the prime-probe task, wherein an initial distractor (or prime) precedes a subsequent target (or probe). In each trial of this task, the distractor may indicate the same response as the target (congruent trials) or a different response (incongruent trials). As one might expect, average reaction time is generally slower in incongruent trials than in congruent trials. Interestingly, however, this congruency effect is often reduced after incongruent (vs. congruent) (Gratton, Coles, & Donchin, 1992; Kunde & Wuhr, 2006).

In general, this *congruency sequence effect* (CSE) is thought to reflect cognitive control processes that minimize distraction and adapt to recent events (Botvinick, Braver, Barch, Carter, & Cohen, 2001; Gratton et al., 1992). Yet, alternative views exist in the literature. According to learning and memory accounts of the CSE, for example, the CSE reflects feature integration (e.g., stimulus repetition) and/or contingency learning (e.g., stimulus frequency) processes that are often confounded with trial congruency (Hommel, Proctor, & Vu, 2004; Mayr, Awh, & Laurey, 2003; Schmidt & De Houwer, 2011). Contrary to these views, however, researchers have successfully developed confound-minimized paradigms that allow one to observe a CSE even in the absence of such confounds (Schmidt & Weissman, 2014; Weissman, Egner, Hawks, & Link, 2015; Weissman, Jiang, & Egner, 2014).

Within the literature, a primary view of this *control-driven* CSE is the episodic retrieval account (Egner, 2014; Hazeltine, Lightman, Schwarb, & Schumacher, 2011; Spapé & Hommel, 2008; Spapé & Hommel, 2014). In this view, various trial features (e.g., stimuli and responses, task representations, etc.) are stored in the memory of the previous trial. During task performance, repeating any of these features in the current trial cues the episodic retrieval of the other previous-trial features – including the control processes underlying the CSE. Critically, this retrieval process biases control processes to prepare for an upcoming trial that resembles the previous one, leading to a reduced congruency effect after incongruent trials than after congruent trials (i.e., a CSE).

Various findings support the episodic retrieval account. For instance, repeating (vs. alternating) stimuli and/or responses across consecutive trials amplifies the CSE (Hommel et al., 2004; Mayr et al., 2003; Weissman et al., 2016). Moreover, repeating context-defining trial features such as stimulus format (Dignath, Johannsen, Hommel, & Kiesel, 2019), sensory modality (Hazeltine, Lightman, Schwarb, & Schumacher, 2011, Kreutzfeldt, Stephan, Willmes, & Koch, 2016), and voice gender (Spapé & Hommel, 2008) amplifies the CSE, whereas alternating these features eliminates or reduces this effect. Similarly, in task-switching paradigms repeating (vs. alternating) the entire task amplifies the CSE even when both tasks involve the same stimuli (Kiesel, Kunde, & Hoffmann, 2006).

However, there is disagreement regarding certain aspects of the episodic retrieval view. More specifically, there are various (and often competing) explanations of these boundary conditions for the CSE. To differentiate among the various explanations, the following studies investigated the boundaries of adaptive control. I briefly describe each of these studies in the sections below.

In the first study (Chapter 2), we investigated whether, and under what conditions, the control processes underlying the CSE minimize cross-modal distraction. The contingent attentional capture hypothesis posits that a cross-modal CSE should occur only when a distractor possesses target-defining features (Moore & Weissman, 2010; Serences et al., 2005; Schmidt & Weissman, 2015). In contrast, the perceptual conflict hypothesis posits that a cross-modal CSE should occur only when perceptual conflict is present between a distractor and target (Notebaert & Verguts, 2006). To distinguish between these two hypotheses, we conducted a pair of experiments involving a cross-modal prime-probe task wherein an auditory distracter word preceded a visual target that appeared in one of two formats (i.e., word or arrow). Consistent with the contingent attentional capture hypothesis, but inconsistent with the perceptual conflict hypothesis, we observed cross-modal CSEs only when the distractor possessed target-defining features. These findings suggest that task-relevant, target-defining features greatly influence the boundaries of adaptive control. This study was published by Acta Psychologica in 2017 with Daniel Weissman.

In the second study (Chapter 3), we investigated why the size of the CSE is at its maximum when both the prime and the probe are task-relevant in a small percentage of trials (Weissman, Colter, Grant, & Bissett, 2017). Given the rare occurrence of such trials, we investigated whether this maximal CSE occurs because the appearance of an occasional task-relevant prime (a) constitutes a rare, unexpected event that triggers heightened control (e.g., Wessel & Aron, 2017) or (b) allows participants to use the same task set for both the prime and probe (e.g., Hazeltine et al., 2011; Spapé & Hommel, 2008). We observed a relatively large CSE even when these "catch" trials occurred in half the trials (Experiment 1) while ruling out an alternative explanation (Experiment 2). Thus, consistent with the latter hypothesis, these findings provide novel support for the idea that *task sets* greatly influence the boundaries of adaptive control. This study was published by Acta Psychologica in 2019 with Daniel Weissman.

In the third study (Chapter 4), we investigated whether task sets or episodic context determine the boundaries of the CSE. To do so, we revisited an ongoing controversy over whether contextual boundaries of the CSE reflect an attentional reset that occurs whenever a context-defining feature changes (attentional reset hypothesis) or only when this change in context-defining features is linked to a different task set (task set hypothesis). To distinguish between these two hypotheses, we employed a cross-modal prime-probe task wherein the sensory modality in which stimuli appear (auditory or visual) varied across trials. Consistent with the task set hypothesis, but not with the attentional reset hypothesis, changing the sensory modality across consecutive trials eliminated the CSE only when the task structure allowed participants to employ this feature to form modality-specific task sets (Experiments 1 and 2). Further, such "modality-specific" CSEs were associated with orienting attention to the sensory modality in which task stimuli appear, which may facilitate task set formation (Experiment 3). Together, these findings suggest that task sets determine the boundaries of adaptive control. This study was published by the Journal of Experimental Psychology: Human Perception and Performance with Savannah Cookson and Daniel Weissman.

In the fourth study (Chapter 5), we investigated whether context-defining features or task sets based on such features determine the boundaries of the CSE in standard, unimodal prime-probe tasks. To distinguish between these two hypotheses, we manipulated whether the structure of a standard, all-visual prime-probe task allowed (Experiment 1) or did not allow (Experiment 2) participants to associate visual context-defining features (i.e., color patches and color words) with different task sets. Consistent with the task set hypothesis, we found that changing salient visual features across consecutive trials reduced the CSE only when the task structure allowed participants to link these visual features to different task sets. These findings extend prior support for the task set hypothesis from somewhat atypical cross-modal tasks to a standard unimodal task. In contrast, they pose a challenge to the attentional reset hypothesis, and other related views, wherein changing salient perceptual features always results in a boundary for the CSE. This study was submitted to the Quarterly Journal of Experimental Psychology with Samantha Cerpa and Daniel Weissman.

In the fifth study (Chapter 6), we investigated the task set hypothesis in the context of episodic retrieval and action control. Currently, two primary views posit that forming and retrieving memories of ongoing events influences action control. However, the organizational structure of these memories, or event files, remains unclear. The hier*a*rchical coding view posits a hierarchical structure, wherein task sets occupy a high level of the hierarchy (Hazeltine et al., 2011; Schumacher & Hazeltine, 2016; Cookson, Hazeltine, & Schumacher, 2016, 2019). Here, the contents of an event file can be retrieved only if the task set repeats. In contrast, the binary coding view posits a non-hierarchical structure, which consists of a collection of independent, binary bindings between different feature pairs (Hommel, 1998; Dignath et al., 2019; Frings et al., 2020). In this view, repeating an abstract feature from a previous event (e.g., the previous trial's S-R mapping) triggers the retrieval of the associated feature from the same binding (e.g., the previous trial's congruency) even if the task set changes. To distinguish between these views, we investigated whether or not a CSE appears when the task set changes but the previous trial's S-R mapping repeats. Two experiments involving a cross-modal prime-probe task yielded a CSE under these conditions while ruling out alternative explanations. These findings show that the typical binary structure of event files generalizes from concrete features (e.g., colors and locations) to abstract features (e.g., S-R mappings and task sets). Moreover, contrary to the hierarchical coding view, they provide a non-hierarchical explanation of task set boundaries for the CSE. This study was submitted to the Journal of Experimental Psychology: Learning, Memory, and Cognition with Daniel Weissman.

Chapter 2 An Attentional Mechanism for Minimizing Cross-Modal Distraction

Abstract

Prior findings suggest that coping with distraction relies on cognitive control processes that increase attention to task-relevant processing, decrease attention to task-irrelevant processing, or both. Consistent with this view, the congruency effect in unimodal Stroop-like tasks, a measure of distraction, is often smaller after more distracting incongruent trials than after less distracting congruent trials. It remains unclear, however, whether, and under what conditions, the control processes underlying this congruency sequence effect (CSE) minimize cross-modal distraction. The contingent attentional capture hypothesis predicts a cross-modal CSE when a distracter possesses a target-defining feature. In contrast, the perceptual conflict hypothesis predicts a cross-modal CSE when there is perceptual conflict between a distracter and a target. To distinguish between these hypotheses, we conducted two experiments wherein an auditory distracter word preceded a visual target that appeared in one of two formats (i.e., word or arrow). We observed robust, cross-modal CSEs. Moreover, the pattern of CSEs that we observed was more consistent with the contingent attentional capture hypothesis than with the perceptual conflict hypothesis. These findings reveal a novel attentional mechanism for minimizing cross-modal distraction.

Introduction

The ability to minimize distraction is crucial for completing everyday tasks. For example, consider an individual who is reading by mentally pronouncing each word. This individual may become distracted by two people speaking nearby, especially if those people say the words he or she is reading. Therefore, to read successfully, this individual must minimize distraction from the irrelevant conversation.

Laboratory studies of distraction

In the laboratory, researchers investigate distraction using distracter interference tasks such as the flanker, Stroop, Simon, and prime-probe tasks (B. A. Eriksen & Eriksen, 1974; Stroop, 1935; Simon & Rudell, 1967; C. W. Eriksen & Schultz, 1979). In each trial of such tasks, participants respond to a target while ignoring a distracter. For example, in the "directional word" version of the prime-probe task, participants indicate the direction specified by a target word (e.g., left, right, up, or down) while ignoring a preceding distracter word (e.g., left, right, up, or down) (Schmidt & Weissman, 2014). In congruent trials, the distracter and target indicate the same direction and thus engender the same response. In incongruent trials, they indicate different directions and thus engender different responses. Typically, participants are slower to respond in incongruent than in congruent trials. This phenomenon, which indexes distractibility, is called the *congruency effect*.

Interestingly, the congruency effect is smaller when the previous trial was incongruent than when it was congruent. This *congruency sequence effect* (CSE) is thought to index a mixture of top-down attentional control processes that minimize distraction (e.g., Gratton, Coles,

& Donchin, 1992; Botvinick, Braver, Barch, Carter, & Cohen, 2001; Kerns, Cohen, MacDonald III, Cho, Stenger, & Carter, 2004) and feature integration and contingency learning processes that are often confounded with trial congruency (Hommel, Proctor, & Vu, 2004; Mayr, Awh, & Laurey, 2003; Schmidt & De Houwer, 2011). However, the use of confound-minimized paradigms has revealed robust CSEs without such confounds (Schmidt & Weissman, 2014; Weissman, Egner, Hawks, & Link, 2015; Weissman, Jiang, & Egner, 2014). Researchers may therefore use these paradigms to investigate the "control-driven" component of the CSE.

A role for contingent attentional capture in triggering the CSE

Recent findings indicate that contingent attentional capture plays a key role in triggering the control-driven CSE. Contingent attentional capture is a phenomenon in which distracters that possess target-defining perceptual and/or categorical features involuntarily capture attention (Moore & Weissman, 2010; Serences et al., 2005). For instance, when participants are asked to identify red targets that appear at a pre-specified location, they are more highly distracted by red distracters that appear at a different location than by other-colored (e.g., blue) distracters (Folk, Leber, & Egeth, 2002). Such capture is thought to occur because participants maintain perceptual and/or categorical representations of target-defining features in working memory. While the maintenance of such an *attentional set* effectively guides attention toward targets, it also permits irrelevant stimuli that possess target-defining features to attract, or capture, attention.

Given that distracters often possess target-defining features in studies of the CSE, Schmidt & Weissman (2015) hypothesized that contingent attentional capture plays a role in triggering this phenomenon. To test this hypothesis, the authors conducted two experiments. In each experiment, participants identified the direction specified by a target (left, right, up, or down) while ignoring the direction specified by a preceding distracter (left, right, up, or down). Critically, the authors varied whether a distracter possessed a target-defining feature. As we describe next, the findings from these experiments indicated an important role for contingent attentional capture in triggering the CSE.

In Experiment 1, the authors varied both the distracter's presentation format (word or arrow) and the target's presentation format (word or arrow) across four blocks. In two blocks, these stimuli appeared in the same format (i.e., both words or both arrows). In the other two blocks, they appeared in different formats (i.e., one word and one arrow). Thus, in each block, participants could adopt an attentional set for just one target format (e.g., words). The authors hypothesized that the distracter would possess a target-defining format, and thereby capture attention, only when it appeared in the same format as the target. Consistent with their hypothesis, the authors observed a larger CSE in blocks wherein the distracter and target appeared in the same format than in blocks wherein these stimuli appeared in different formats. In fact, the CSE was completely absent in the latter condition. These findings provided novel support for the view that contingent attentional capture plays a role in triggering the CSE.

In Experiment 2, the authors randomly varied the format of the distracter and target on a trial-by-trial basis. Since the target in each trial could be either a word or an arrow, participants had to adopt an attentional set for both target formats. The authors therefore hypothesized that a distracter in either format would possess a target-defining feature that captures attention. Consistent with their hypothesis, and unlike in Experiment 1, the authors observed equivalent, robust CSEs in the same and different format conditions. These findings provided further support for the view that contingent attentional capture triggers the CSE.

Does contingent attentional capture trigger a cross-modal CSE?

As in the example of reading described earlier, irrelevant distracters often appear in a different sensory modality than relevant targets. To our knowledge, however, researchers have yet to investigate whether control processes underlying the CSE minimize cross-modal distraction. Investigating this hypothesis is important for three reasons. First, multisensory interactions are linked to unique distraction effects (Zimmer, Roberts, Harshbarger, Woldorff, 2010). Second, some methods for reducing unimodal distraction (e.g., increasing the perceptual load of a visual display) do not minimize cross-modal distraction (Tellinghuisen & Nowak, 2003). Third, such a finding would expand the domain over which control processes underlying the CSE are known to operate.

Based on our prior findings (Schmidt & Weissman, 2015), we hypothesize that the CSE will minimize cross-modal distraction when a distracter possesses a target-defining feature that captures attention. Along these lines, words that appear in the auditory modality share perceptual and/or categorical features (i.e., auditory-verbal representations) with words that appear in the visual modality. For example, recent findings from functional magnetic resonance imaging (fMRI) indicate that auditory voice-selective regions are activated during silent reading (Yao, Belin, and Scheepers, 2011). These findings suggest that the attentional set people adopt to read visual words includes auditory-verbal representations of those words. Thus, a word spoken in headphones (e.g., “Left”) should capture attention if it is a potential target that can appear on a computer screen (e.g., “Left”). The *contingent attentional capture* hypothesis therefore predicts that auditory distracter words will trigger a CSE when participants adopt an attentional set that specifies analogous visual target words.

Interestingly, this hypothesis further predicts that auditory distracter words will not capture attention when participants can adopt a non-verbal attentional set for visual targets. For example, according to this hypothesis, auditory distracter words should not trigger a control-driven CSE when the visual target is always an arrow, because task-relevant arrows are represented with spatial, rather than with auditory-verbal, features (Miles & Proctor, 2011). In this situation, the attentional set that participants adopt to search for target arrows should not include auditory-verbal representations that are activated by auditory distracter words. Thus, auditory distracter words should not trigger a CSE when the visual target is always an arrow.

Does perceptual conflict trigger a cross-modal CSE?

Contrary to the contingent attentional capture hypothesis, the *perceptual conflict* hypothesis posits that the CSE is triggered by perceptual conflict between the distracter and the target in incongruent trials (Notebaert & Verguts, 2006). For instance, perceptual conflict may occur when an auditory distracter word and a visual target word differ, because these stimuli activate distinct auditory-verbal representations. However, such conflict is unlikely to occur between an auditory distracter word and a visual target arrow, because the auditory-verbal representations that are employed to identify an auditory word are unlikely to conflict at perceptual levels with the spatial representations that are employed to identify a visual target arrow. Thus, according to the perceptual conflict hypothesis, a cross-modal CSE should be observed when an auditory distracter and a visual target are both words, but not when the distracter is a word and the target is an arrow. While findings from unimodal visual-modality tasks weigh against the perceptual conflict hypothesis (Schmidt & Weissman, 2015), to our knowledge no prior work has investigated this hypothesis in a cross-modal context.

The Present Study

The goal of the present study was to investigate, and ultimately distinguish between, the contingent attentional capture and perceptual conflict hypotheses. To this end, we conducted two experiments. Both experiments were similar in design to those of Schmidt & Weissman (2015). In each experiment, however, the distracter and target appeared in different sensory modalities. The distracter was an auditory word (“Left”, “Right”, “Up”, or “Down”). The target was either a visual word (Left, Right, Up, or Down) or a visual arrow that pointed left, right, up, or down.

In Experiment 1, we varied the target format (word or arrow) across blocks. Thus, in each block, participants could adopt an attentional set for just one target format (word *or* arrow). In this situation, auditory distracter words should possess target-defining perceptual or categorical features when the visual targets are analogous words but not when they are arrows. Similarly, perceptual conflict between the auditory distracter and the visual target should be greater in incongruent than in congruent trials when both stimuli are words, but not when the distracter is a word and the target is an arrow. Thus, both hypotheses predict the CSE will be larger in word-target trials than in arrow-target trials. The goal of Experiment 1 was to test this hypothesis and thereby establish whether a cross-modal CSE is observed only when auditory distracters activate the same auditory-verbal representations as potential visual targets. Such an outcome would reveal an important boundary condition for observing a cross-modal CSE.

In Experiment 2, we randomly varied the target format (word or arrow) on a trial-by-trial basis. Thus, in each block, participants were required to adopt an attentional set for both target formats (word *and* arrow). In this situation, auditory distracter words should always possess target-defining perceptual or categorical features. The contingent attentional capture hypothesis therefore predicts equivalent, robust CSEs in both word- and arrow-target trials. In contrast, the perceptual conflict hypothesis still predicts a CSE in word-target trials but not in arrow-target trials. Indeed, an auditory distracter word is unlikely to conflict with a visual target arrow at perceptual levels.

Experiment 1

In Experiment 1, we investigated whether a cross-modal CSE is observed only when auditory distracters activate the same auditory-verbal representations as potential visual targets. To do so, we asked participants to perform a modified version of the prime-probe task employed by Schmidt and Weissman (2015). In each trial, an auditory distracter word ("Left", "Right", "Up", or "Down") preceded a visual target, which was a word in some blocks and an arrow in others. Participants were asked to indicate the direction specified by the target (Left, Right, Up, or Down) while ignoring the distracter. Consistent with both the contingent attentional capture and perceptual conflict hypotheses, we predicted that the CSE would be larger in word-target trials than in arrow-target trials.

Methods

Participants

Thirty-six undergraduates at the University of Michigan participated in Experiment 1 in exchange for course credit. We excluded the data from four participants due to a script error (one participant), self-reported neurological disorders and/or psychoactive medications (two participants), and performing the task with less than 75% accuracy (one participant). None of the remaining participants (12 male, 20 female; 18-20 years old; mean age, 18.5 years; SD = 0.72

years) reported any history of head trauma, uncorrected vision or hearing impairments, seizures, or neurological disorders.

Stimuli

Three types of stimuli appeared in each block. These included (1) a central fixation cross (0.8°x 0.8°), (2) an auditory distracter word spoken in headphones (“Left”: 63 dB, “Right”: 64 dB, “Up”: 64 dB, or “Down”: 68 dB), which was created by a MATLAB script that converts written text to speech (https://www.mathworks.com/matlabcentral/fileexchange/18091-text-to-speech), and (3) a central visual target, which was either a word (Left: 4.3° x 1.0°, Right: 5.4° x 1.5°, Up: 2.1° x 1.8°, or Down: 4.3° x 1.0°) or an arrow (Left, Right, Up, or Down; horizontal orientation, 5.5° x 3.2°; vertical orientation, 3.2° x 5.5°).

There were two sets of distracter-target pairs. Each set consisted of four congruent distracter-target pairs (left-left, right-right, up-up, down-down) and four incongruent distracter-target pairs (left-right, right-left, up-down, down-up). Four distracter-target pairs consisted of left and right auditory distracter words paired with left and right visual targets (the left-right set). The remaining four distracter-target pairs consisted of up and down auditory distracter words paired with up and down visual targets (the up-down set). We used the Psychophysics Toolbox to present stimuli and record participants’ responses (Brainard, 1997).

Task and Design

Each block started and ended with a 2000 ms fixation cross. In the interim, there was a sequence of trials. There were four sequential events in each 2.5-second trial: an auditory distracter word (duration, 200-300 ms), a blank screen (duration, 333-433 ms), a target word or arrow (duration, 133 ms), and a second blank screen (duration, 1733 ms)[1]. Participants responded to left words and left arrows by pressing the “F” key with the left middle finger, to right words and right arrows by pressing the “G” key with the left index finger, to up words and up arrows by pressing the “J” key with the right middle finger, and to down words and down arrows by pressing the “N” key with the right index finger. We instructed participants to respond as quickly and accurately as possible to the target in each trial. If participants responded incorrectly or did not respond within 1500 ms of target onset, the word “Error” appeared at the center of the screen for 200 ms.

The experiment was divided into two halves. In one half, the target was an arrow. In the other half, the target was a word (Fig. 1). Each half consisted of a single 32-trial practice block and four 96-trial test blocks. We counterbalanced across participants which half (word or arrow) was presented first.

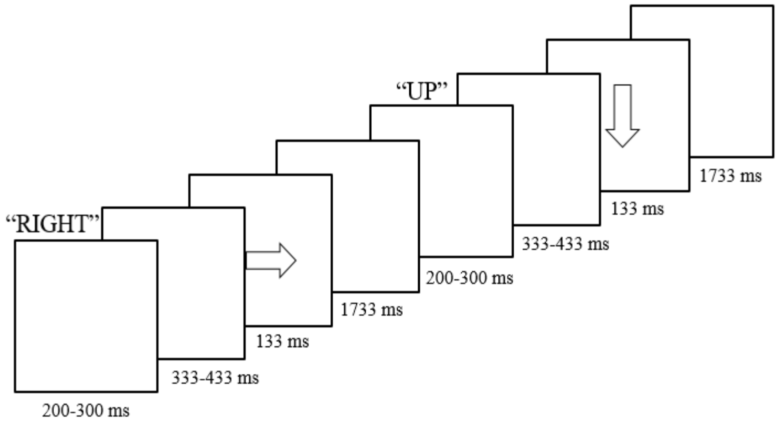


Figure 2-1. The prime-probe task used in Experiment 1

In each trial, an auditory distracter word (shown in quotes) preceded a visual target (two trials are illustrated in the figure). In half of the trial blocks, the target was an arrow as shown above. In the other half of the trial blocks, the target was a word. In the actual task, the target stimuli appeared in white on a black background. The numbers beneath each box indicate the length of each trial component in milliseconds (ms).

---

[1] The duration of the auditory distracter word (approximately 200-300 ms) varied with its identity (i.e., left, right, up, down). Therefore, we adjusted the duration of the subsequent blank screen such that the summed duration of the auditory distracter word and subsequent blank screen was always 633 ms.

As in our prior work (Weissman et al., 2015; Schmidt & Weissman, 2014), we avoided two prominent confounds that are present in many studies of the CSE. First, we avoided feature integration confounds, which occur when there are repetitions of stimulus and/or response features in consecutive trials (Hommel et al., 2004; Mayr et al., 2003). To avoid repetitions, we presented stimuli from the left-right set in odd trials and stimuli from the up-down set in even trials. Second, we avoided contingency learning biases, which occur when a distracter is more frequently associated with a congruent target than with an incongruent target (Schmidt & DeHouwer, 2011). To do so, we presented each of the eight distracter-target pairs equally often in every block.

In each 96-trial block, we first-order counterbalanced the order in which congruent and incongruent trials appeared. Therefore, each block had approximately 24 congruent trials preceded by congruent trials (cC trials), 24 incongruent trials preceded by congruent trials (cI trials), 24 congruent trials preceded by incongruent trials (iC trials), and 24 incongruent trials preceded by incongruent trials (iI trials). The first trial in each block was not preceded by a trial. Thus, one of the four congruency sequences above had only 23 trials. The underrepresented congruency sequence varied randomly across blocks because we used a different counterbalanced trial order for each block of trials.

Procedure

After arriving at the laboratory, the participant provided informed written consent and filled out a brief screening form. Next, the researcher brought the participant to an individual testing chamber. The participant was seated with his or her head in a chinrest to prevent movement during the experiment with the exception of short rest breaks between blocks. The distance from the computer screen to the participant's eyes was 55 cm.

The experiment began with onscreen instructions for the first condition (word targets or arrow targets, counterbalanced across participants). These instructions were also explained verbally to the participant. Next, the participant began the practice block. The researcher remained in the testing chamber to verify that the participant was able to complete task with a high degree of accuracy. If numerous errors were made, the researcher re-explained the instructions and started a second practice block. Once the participant completed the practice block with relatively few errors, the researcher left the chamber. The participant then completed four test blocks with either a word target or an arrow target.

After the four test blocks were completed, the participant informed the researcher that he or she was ready for the second half of the experiment. The researcher then explained the instructions for the second half of the experiment, which were virtually identical to those for the first half. The only difference was that the target appeared in whichever presentation format (word or arrow) was not presented in the first half of the experiment. As described earlier, the participant completed one practice block and four test blocks.

Data Analysis

We excluded several types of trials in the data analyses. When analyzing mean response time (RT), we excluded practice trials, trials with omitted or incorrect responses, trials following omitted and incorrect responses, and outliers. We defined outliers as trials with RTs that were at least three standard deviations away from their conditional mean. When analyzing mean error rate (ER), we excluded practice trials, trials following an omitted or incorrect response, and outliers. On average, 1.40% of the trials were outliers and 4.24% were errors.

We conducted separate repeated measures analyses of variance (ANOVA) on mean RT and mean ER. Each ANOVA had three within-participants factors: target type (word, arrow), previous congruency (congruent, incongruent), and current congruency (congruent, incongruent). Tables 1 and 2 present the mean RT and mean ER data, respectively, in the main experimental conditions of Experiments 1 and 2.

Results

Mean RT

There were three significant main effects. First, there was a main effect of target type, $F(1,31) = 144.226$, $p < 0.001$, $\eta_p^2 = 0.823$, because mean RT was slower when the target was a word (517 ms) than when it was an arrow (425 ms). Second, there was a main effect of previous congruency, $F(1,31) = 22.982$, $p < 0.001$, $\eta_p^2 = 0.426$, because mean RT was slower following incongruent (474 ms) as compared to congruent (468 ms) trials. Third, there was a main effect of current congruency, $F(1,31) = 163.511$, $p < 0.001$, $\eta_p^2 = 0.841$, because mean RT was slower in incongruent (495 ms) as compared to congruent (446 ms) trials.

There were also two significant two-way interactions. First, there was an interaction between previous congruency and current congruency, $F(1,31) = 12.717$, $p < 0.005$, $\eta_p^2 = 0.291$, because the congruency effect was smaller after incongruent (44 ms) as compared to congruent (54 ms) trials. Second, there was an interaction between target type and current congruency, $F(1,31) = 13.510$, $p = 0.001$, $\eta_p^2 = 0.304$, because the congruency effect was smaller when the target was an arrow (40 ms) than when it was a word (57 ms).

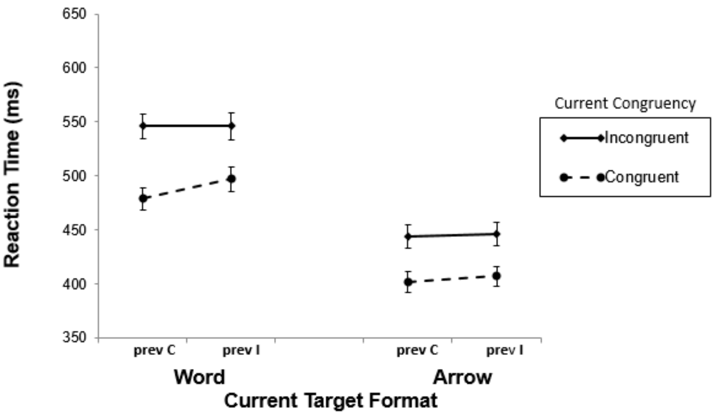


Figure 2-2. The main findings from Experiment 1 The CSE was larger when the target was a word (left) than when the target was an arrow (right). On the x-axis, "prev C" and "prev I" indicate that the previous trial was congruent or incongruent. Error bars indicate the standard error of the mean.

Finally, there was a significant three-way interaction. Specifically, there was an interaction among target type, previous congruency, and current congruency, $F(1,31) = 5.330$, $p < 0.05$, $\boxed{\eta_p^2} = 0.147$ (Fig. 2). Consistent with both hypotheses, this interaction occu word (18 ms; $F(1,31) = 14.382$, $p = 0.001$, 0.563, $\boxed{\eta_p^2} = 0.018$). No other effects were significant.

Table 2-1. Mean RT in Experiments 1 and 2

	cC	**cI**	**iC**	**iI**	**Mean**	**I – C**	**CSE**
Experiment 1							
Word-target trials	479	546	497	546	517	57	18
	(10)	(12)	(12)	(12)	(13)	(5)	(5)
Arrow-target trials	402	444	407	446	425	40	3
	(10)	(10)	(10)	(11)	(11)	(4)	(4)
Experiment 2							
Word-target trials	512	584	527	581	551	63	18
	(13)	(16)	(13)	(16)	(15)	(7)	(5)
Arrow-target trials	449	507	463	504	481	49	16
	(12)	(16)	(13)	(16)	(15)	(7)	(4)

Note. cC = previous trial congruent, current trial congruent; cI = previous trial congruent, current trial incongruent; iC = previous trial incongruent, current trial congruent; iI = previous trial incongruent, current trial incongruent; I-C = incongruent– congruent; CSE = congruency sequence effect, calculated as: (cI - cC) – (iI - iC).

Mean ER

There were two significant main effects. First, there was a main effect of target type,

$F(1,31) = 20.712$, $p < 0.001$, $\eta_p^2 = 0.401$, because mean ER was greater when the target was a word (4.4%) than when it was an arrow (2.8%). Second, there was a main effect of current congruency, $F(1,31) = 33.948$, $p < 0.001$, $\eta_p^2 = 0.523$, because mean ER was greater in incongruent (4.7%) as compared to congruent (2.4%) trials.

Table 2-2. Mean error rate in Experiments 1 and 2

	cC	**cI**	**iC**	**iI**	**Mean**	**I – C**	**CSE**
Experiment 1							
Word-target trials	2.7	6.5	3.1	5.3	4.4	3.0	1.6
	(0.4)	(0.8)	(0.5)	(0.6)	(0.6)	(0.4)	(0.9)
Arrow-target trials	2.2	3.7	1.7	3.5	2.8	1.6	-0.2
	(0.4)	(0.7)	(0.3)	(0.6)	(0.5)	(0.5)	(0.6)
Experiment 2							
Word-target trials	3.6	5.3	2.9	5.1	4.2	1.9	-0.4
	(0.8)	(0.7)	(0.5)	(0.8)	(0.7)	(0.5)	(0.8)
Arrow-target trials	1.0	2.9	1.0	2.2	1.8	1.5	0.7
	(0.2)	(0.6)	(0.3)	(0.4)	(0.4)	(0.4)	(0.6)

Note. cC = previous trial congruent, current trial congruent; cI = previous trial congruent, current trial incongruent; iC = previous trial incongruent, current trial congruent; iI = previous trial incongruent, current trial incongruent; I-C = incongruent– congruent; CSE = congruency sequence effect, calculated as: (cI - cC) – (iI - iC).

Discussion

In Experiment 1, we investigated whether control processes underlying the CSE minimize cross-modal distraction when a distracter appears in a target-defining presentation format (Schmidt & Weissman 2015). Consistent with this possibility, we found that auditory distracter words triggered a cross-modal CSE when the visual targets were corresponding directional words, but not when they were corresponding directional arrows. This finding is consistent with the contingent attentional capture hypothesis. In particular, it fits with prior work indicating that participants include auditory-verbal representations in the attentional set they employ to search for visual word targets, but not in the attentional set they employ to search for visual arrow targets (Miles & Proctor, 2011). This finding is also consistent with the perceptual conflict hypothesis. More specifically, it fits with the view that perceptual conflict is more likely when distracters and targets activate potentially conflicting perceptual representations (e.g., representations of visual and spoken words) than when they activate qualitatively different perceptual representations (e.g., representations of words and arrows). In short, our findings are consistent with both the contingent attentional capture and perceptual conflict hypotheses.

Experiment 2

The goal of Experiment 2 was to distinguish between the contingent attentional capture and perceptual conflict hypotheses. To this end, we investigated whether a CSE is observed in both word- and arrow-target trials when these trials are randomly intermixed in the same blocks. In this situation, participants must adopt an attentional set for both word and arrow presentation formats throughout each block of trials. An auditory distracter word should therefore possess a target-defining feature, and thereby capture attention, regardless of whether a visual target is a word or an arrow. The contingent attentional capture hypothesis therefore predicts equivalent, robust CSEs in word- and arrow-target trials. In contrast, the perceptual conflict account still predicts the absence of a CSE in arrow-target trials. Indeed, the auditory-verbal representations that participants use to identify an auditory distracter word cannot conflict at perceptual levels with the spatial representations that participants use to identify a visual target arrow. Perceptual conflict should therefore be absent in arrow-target trials.

Methods

Participants

Thirty-nine undergraduates at the University of Michigan received course credit for participating in Experiment 2. We excluded the data from seven participants for not completing the experiment (one participant), self-reported neurological disorders and/or psychoactive medications (two participants), technical issues (one participant), and performing the task with less than 75% accuracy (three participants). None of the remaining participants (14 male, 18 female; 18-20 years old; mean age, 18.81 years; SD = 0.72) reported any history of head trauma, uncorrected vision or hearing impairments, seizures, or neurological disorders.

Stimuli

The stimuli were identical to those in Experiment 1.

Task and Design

The task and design were identical to those in Experiment 1 with one exception. Namely, rather than presenting the word and arrow targets in different blocks, we randomly intermixed these targets in the same blocks (Fig. 3). Thus, the experiment consisted of a single 32-trial practice block followed by eight 96-trial test blocks.

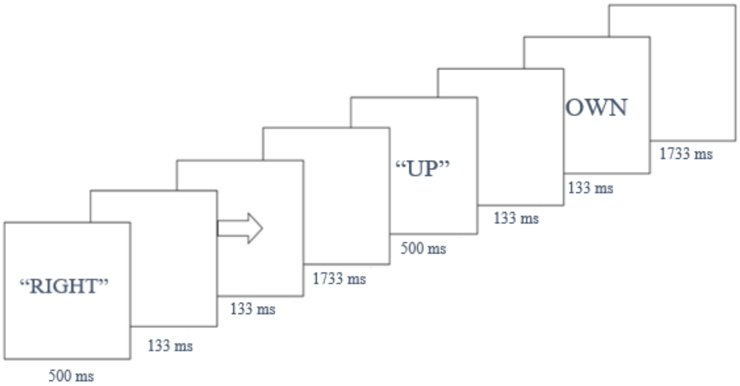


Figure 2-3. The prime-probe task used in Experiment 2
In each trial, an auditory distracter word (shown in quotes) preceded a visual target (two trials are illustrated in the figure). Unlike in Experiment 1, the target format (word or arrow) varied on a trial-by-trial basis within each block. In the actual task, the target stimuli appeared in white on a black background. The numbers beneath each box indicate the length of each trial component in milliseconds (ms).

In each test block, we first-order counterbalanced the sequence in which four trial types appeared: congruent word targets, incongruent word targets, congruent arrow targets, and incongruent arrow targets. We did so by requiring that each trial type be preceded approximately equally often by all four trial types in the experimental design. Thus, in each block, there were approximately six trials in each of the 16 possible combinations of previous target type (word, arrow), current trial type (word, arrow), previous congruency (congruent, incongruent), and current congruency (congruent, incongruent). For example, when a word target appeared in both the previous and the current trial, each block had approximately six congruent trials preceded by congruent trials (cC trials), six incongruent trials preceded by congruent trials (cI trials), six congruent trials preceded by incongruent trials (iC trials), and six incongruent trials preceded by incongruent trials (iI trials). Since the first trial in each block was not preceded by a trial, one of the 16 possible combinations mentioned above appeared only five times in a block. The underrepresented congruency sequence varied randomly across blocks, because we used a different counterbalanced trial order in every block.

Procedure

The procedure was identical to that in Experiment 1 with one exception: participants performed the same task in all eight test blocks, rather than one task in four blocks and a slightly different task in the remaining four blocks. Given this change, participants did not notify the researcher that they were ready for the next task midway through the experiment.

Data Analysis

We excluded the same trial types as in Experiment 1. On average, 0.32% of the trials were outliers and 3.61% were errors. Finally, we analyzed the mean RT and mean ER data in separate repeated-measures ANOVAs with four factors: previous target type (word, arrow), current target type (word, arrow), previous congruency (congruent, incongruent), and current congruency (congruent, incongruent).

Results

Mean RT

There were three significant main effects. First, there was a main effect of current target type, $F(1,31) = 269.814$, $p < 0.001$, $\eta_p^2 = 0.897$, because mean RT was slower when the current target was a word (551 ms) than when it was an arrow (481 ms). Second, there was a main effect of previous congruency, $F(1,31) = 11.604$, $p < 0.005$, $\eta_p^2 = 0.272$, because mean RT was slower following incongruent (519 ms) as compared to congruent (513 ms) trials. Third, there was a main effect of current congruency, $F(1,31) = 66.827$, $p < 0.001$, $\eta_p^2 = 0.6833$, because mean RT was slower in incongruent (544 ms) than in congruent (488 ms) trials.

There were also three significant two-way interactions. First, there was an interaction between previous congruency and current congruency, $F(1,31) = 24.951$, $p < 0.001$, $\eta_p^2 = 0.446$, because the congruency effect was smaller after incongruent (48 ms) as compared to congruent (64 ms) trials. Second, there was an interaction between previous target type and current target type, $F(1,31) = 38.214$, $p < 0.001$, $\eta_p^2 = 0.552$: although participants always responded more slowly to word targets than to arrow targets, this effect was larger when the previous target was an arrow (80) ms relative to a word (60 ms). Third, there was an interaction between current target type and current congruency, $F(1,31) = 14.032$, $p = 0.001$, $\eta_p^2 = 0.312$, because the congruency effect was smaller in arrow (49 ms) than in word (63 ms) trials.

Critically, unlike in Experiment 1, the three-way interaction among current target type, previous congruency, and current congruency was not significant, $F(1,31) = 0.070$, $p = 0.793$, $\eta_p^2 = 0.002$ (Fig. 4). This is because the CSE did not differ between word-target trials (18 ms; $F(1,31) = 12.074$, $p = 0.002$, $\eta_p^2 = 0.280$) and arrow-target trials (16 ms; $F(1,31) = 13.865$, $p = 0.001$, $\eta_p^2 = 0.309$). Finally, the three-way interaction among previous target type, previous congruency, and current congruency was not significant, $F(1,31) = 1.061$, $p = 0.215$, $\eta_p^2 = 0.049$. This is because the magnitude of the CSE did not vary with whether the previous target type was a word (20 ms; $F(1,31) = 16.810$, $p < 0.001$, $\eta_p^2 = 0.352$) or an arrow (13 ms; $F(1,31) = 10.425$, $p < 0.005$, $\eta_p^2 = 0.252$). No other effects were significant.

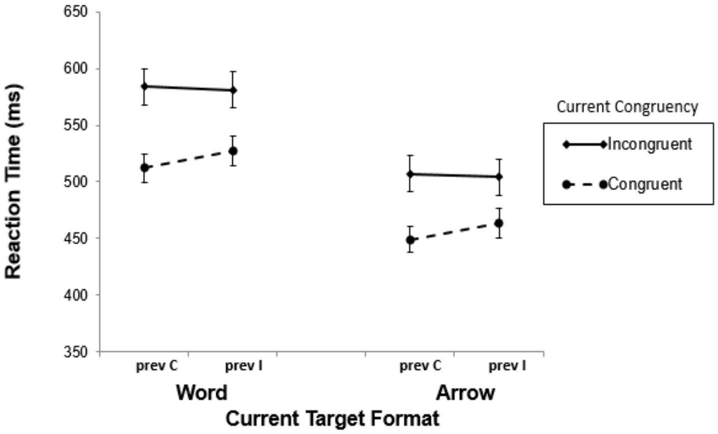


Figure 2-4. The main findings from Experiment 2
We observed equivalent, robust CSEs regardless of whether the current target was a word (left) or an arrow (right). On the x-axis, "prev C" and "prev I" indicate that the previous trial was congruent or incongruent. Error bars indicate the standard error of the mean.

Mean ER

There were two significant main effects. First, there was a main effect of current target type, $F(1,31) = 30.127$, $p < 0.001$, $\eta_p^2 = 0.493$, because mean ER was greater in word (4.2%) as compared to arrow (1.8%) trials. Second, there was a main effect of current congruency, $F(1,31)$

= 20.426, $p < 0.001$, $\boxed{\eta_p^2}$ = 0.397, because mean ER was greater in incongruent (3.9%) as compared to congruent (2.1%) trials.

There was also one significant two-way interaction. Specifically, there was an interaction between previous target type and current congruency, $F(1,31) = 5.375$, $p < 0.05$, $\boxed{\eta_p^2}$ = 0.148, because the congruency effect was smaller when the target in the previous trial was a word (1.2%) as compared to an arrow (2.3%). No other effects were significant.

Discussion

The results of Experiment 2 favor the contingent attentional capture hypothesis over the perceptual conflict hypothesis. In this experiment, word- and arrow-target trials appeared randomly in each block. Participants therefore always needed to adopt an attentional set that specified the auditory-verbal representations of visual words, meaning that the auditory analogs of these words should capture attention in both word- and arrow-target trials. Although the contingent attentional capture and perceptual conflict hypotheses are not necessarily mutually exclusive, the equivalent CSEs that we observed in word- and arrow-target trials are more consistent with the contingent attentional capture hypothesis than with the perceptual conflict hypothesis. Thus, we conclude that our results extend to cross-modal situations prior findings indicating that contingent attentional capture triggers the CSE (Schmidt & Weissman, 2015).

General Discussion

The present findings indicate for the first time that control processes underlying the CSE minimize cross-modal distraction. Moreover, analogous to previous findings from unimodal visual-modality experiments (Schmidt & Weissman, 2015), they reveal that this effect is more consistent with the contingent attentional capture hypothesis than with the perceptual conflict hypothesis. These findings make several important contributions to the literature.

A role for contingent attentional capture in triggering the CSE

The present findings complement prior data from unimodal, visual-modality experiments indicating that contingent attentional capture triggers the CSE (Schmidt & Weissman, 2015). As such, they further suggest that the human brain has evolved attentional mechanisms for minimizing distraction that operate only (or mainly) on irrelevant stimuli that possess target-defining features (Schmidt & Weissman, 2015). While it may appear surprising that an attentional mechanism that minimizes distraction does not operate on all irrelevant stimuli, distracters often disrupt goal-directed behavior to a greater degree when they possess a target-defining feature than when they do not (Folk et al., 2002; Moore & Weissman, 2010; Serences et al., 2005). From this perspective, it is perhaps less surprising that the brain employs specialized, modality-general mechanisms for minimizing distraction from such stimuli.

Our finding that contingent attentional capture triggers cognitive control processes that minimize distraction may also appear surprising given prior work indicating that such capture typically increases distraction. For example, participants often identify visual targets less accurately when they are preceded by target-colored as compared to non-target-colored distracters (Folk et al., 2002; Moore & Weissman, 2010). In our view, these two sets of findings are not incompatible. Rather, they suggest that distracters possessing target-defining features trigger cognitive control processes *because* they capture attention and may undermine performance. Subsequently, these stimuli may lead control processes to update an attentional set for trial congruency, resulting in reduced distracter processing after incongruent relative to congruent trials and, hence, a CSE (Schmidt & Weissman, 2015; Weissman, Colter, Drake, & Morgan, 2015).

Broader relevance of our findings

Since the present findings extend the domain over which control processes underlying the CSE are known to operate to cross-modal task situations, it is interesting to consider whether these control processes might also operate in other domains wherein potent distracters threaten to undermine performance. For example, irrelevant stimuli that possess target-defining, reward-related features (e.g., colors) are often highly distracting (Krebs et al., 2010; Anderson, 2015a; Anderson, 2015b; Anderson, Folk, Garrison, & Rogers, 2016). Thus, future research might investigate whether control processes underlying the CSE also reduce reward-related distraction. Such a finding would inform prior data indicating that control processes often fail to minimize the influence of reward-related stimuli on performance (Krebs et al., 2010; Anderson, 2015a; Anderson, 2015b; Anderson et al., 2016). Specifically, it would reveal a novel attentional mechanism for minimizing this influence. Given that certain disorders of attention (e.g., drug addiction) are linked to heightened distraction from reward-related stimuli (Robinson & Berridge, 2008), confirming this hypothesis could lead to important clinical applications.

The present findings also suggest that the control processes triggered by incongruent trials are more effective at minimizing cross-modal distraction than the control processes that are triggered by increasing perceptual load. According to the *perceptual load* theory of attention, irrelevant distracters are identified only if relevant stimuli do not exhaust perceptual attentional resources (Lavie, 1995; Lavie & Tsal, 1994). Under such “low load” conditions, irrelevant stimuli are identified and engender the standard congruency effect. Under “high load” conditions, however, task-relevant stimuli exhaust perceptual attentional resources. Therefore, irrelevant stimuli are not identified, and the congruency effect is absent. Consistent with this theory, increasing the perceptual load of a visual display minimizes distraction from within-modality visual distracters. (e.g., Lavie, 2005). However, it does not minimize distraction from across-modality auditory distracters (Tellinghuisen & Nowak, 2003). Our findings therefore suggest that the control processes underlying the CSE are more effective at minimizing cross-modal distraction than the control processes underlying perceptual load effects.

Finally, the present findings suggest new research directions that could advance our understanding of real-world cross-modal distraction. For example, they suggest it could be worthwhile to investigate whether either the congruency effect or the CSE in our cross-modal task correlates with real-world measures of auditory distraction. One such measure is the adverse effect of irrelevant spoken words on reading comprehension (e.g., Martin, Wogalter, & Forlano,1988). Since the congruency effect measures overall distractibility while the CSE measures the ability to cope with distraction on a trial-by-trial basis, greater comprehension in the context of irrelevant auditory speech may correlate with smaller congruency effects and/or larger CSEs in our cross-modal task. Future research could investigate these hypotheses.

Implications for theoretical accounts of the control-driven CSE

There are two prominent accounts of the control-driven CSE: the *attentional shift* account and the *response modulation* account. The attentional shift account posits that control processes shift attention toward the target and/or away from the distracter after incongruent relative to congruent trials (Gratton et al., 1992; Botvinick et al., 2001). The response modulation account posits that control processes modulate the response signaled by the distracter before the target response comes online (Ridderinkhof, 2002; Weissman et al., 2015).

With regard to the prime-probe task, recent findings favor the response modulation account over the attentional shift account. For example, when the inter-stimulus-interval (ISI) separating the distracter and the target in a prime-probe task is relatively long (i.e., 1000 ms), which eliminates the overall congruency effect, the CSE is linked to a small positive congruency effect after congruent trials and to a small negative congruency effect (i.e., faster response times in incongruent relative to congruent trials) after incongruent trials (Weissman et al., 2015). A negative congruency effect is not consistent with the attentional shift account. Even shifting all of one's attention to the target would yield no congruency effect, rather than a negative congruency effect. In contrast, a negative congruency effect is consistent with the response modulation account. For example, inhibiting the response signaled by the distracter after an incongruent trial should slow performance more in congruent trials than in incongruent trials and could therefore engender a negative congruency effect in the absence of an overall congruency effect.

We have previously noted that an attentional shift may also contribute to the CSE in prime-probe tasks (Weissman et al., 2014). Most important for present purposes, our cross-modal task could provide a useful tool for investigating this possibility. Unlike in a typical visual-modality task, in our cross-modal task we present the distracter in a different sensory modality and at a different spatial location than the visual target. Thus, relative to a visual-modality task, participants may find it easier to increase attention to the target after an incongruent trial without also increasing attention to the distracter. Such a result would provide novel support for the attentional shift account.

Finally, it is important to note that our findings are specific to the prime-probe task, meaning that different control processes may underlie CSEs in different tasks (Egner, 2008). Consistent with this possibility, in a combined Stroop and Simon task, the CSE that was triggered by stimulus-based (i.e., Stroop) conflict was completely independent of the CSE that was triggered by response-based (i.e., Simon) conflict and vice-versa (Egner, Delano, and Hirsch, 2007). However, these and other "task-specific" CSEs could reflect dissociable associative mechanisms, rather than dissociable control mechanisms, due to the presence of feature integration confounds (Egner, 2008; Hommel, Proctor, and Vu, 2004; Mayr, Awh, and Laurey, 2003). In other words, the task-specific CSEs in these studies could have been driven by task-specific associative mechanisms, rather than by task-specific control processes (Braem, Abrahamse, Duthoo, and Notebaert, 2014). To determine whether the control mechanism underlying the confound-minimized CSE varies across tasks (e.g., response modulation versus attentional shift), future studies could utilize confound-minimized paradigms similar to the one we employed in our cross-modal task.

A potential alternative account of the present findings

Since some variants of the attentional shift account posit that response conflict triggers the CSE, one may wonder whether our findings can be explained by the fact that the congruency effect (i.e., the behavioral signature of response conflict; Yeung et al., 2011) was always larger in word-target trials than in arrow-target trials. We cannot completely rule out an influence of response conflict on the CSE in the present study. We note, however, that while the congruency effect was greater in word-target than in arrow-target trials in both Experiments 1 and 2, the CSE was greater in word-target than in arrow-target trials only in Experiment 1. Thus, we conclude that the pattern of CSEs we have observed is unlikely to be explained by response conflict. Consistent with this view, prior findings from confound-minimized tasks indicate that CSE magnitude does not vary with the size of the congruency effect (Weissman et al., 2015).

Conclusion

The present findings demonstrate for the first time that control processes underlying the CSE minimize cross-modal distraction. Moreover, they indicate that contingent attentional capture plays a key role in triggering this effect. Future research investigating the role of contingent attentional capture in triggering the CSE may shed additional light on how cognitive control processes minimize distraction from irrelevant stimuli.

Chapter 3 Turning Distractors into Targets Increases the Congruency Sequence Effect

Abstract

Congruency effects in distractor-interference tasks are typically smaller after incongruent trials than after congruent trials. Current views posit that this congruency sequence effect (CSE) reflects control processes that minimize distraction from irrelevant stimuli. However, the CSE is counterintuitively larger in the prime-probe task when the prime is occasionally a second target than when the prime is more frequently a distractor. In the present study, we investigated whether this effect occurs because the appearance of an occasional prime target (a) constitutes a rare, unexpected event that triggers heightened control or (b) allows participants to use the same task set (i.e., stimulus-response mapping) for the prime and probe in each trial. Consistent with the latter hypothesis, we observed this effect in Experiment 1 even when the critical trial types appeared equally often. Further, in Experiment 2, we extended this finding while ruling out perceptual differences between conditions as an alternative account. These findings provide novel support for the task set hypothesis and reveal that the CSE reflects control processes that do more than minimize distraction.

Introduction

Minimizing distraction is essential for avoiding predators, navigating environments, and even for playing sports. For instance, consider a basketball player who distracts a defender by looking to the right just before passing the ball to the left. Such “head-fakes” require a defender to minimize distraction from the offensive player’s irrelevant head and gaze orientation.

Laboratory studies of distraction

Researchers often use distracter-interference tasks such as the flanker, Stroop, Simon, and prime-probe tasks to investigate how individuals minimize distraction in the laboratory (Eriksen & Eriksen, 1974; Stroop, 1935; Simon & Rudell, 1967; Eriksen & Shultz, 1979). For instance, consider the prime-probe arrow task. Analogous to a head fake, a distractor arrow (prime) appears just before a target arrow (probe) in each trial. The participant must indicate the direction in which the target arrow points while ignoring the distractor arrow, just as a basketball player must identify the direction in which an opponent passes the ball while ignoring a head fake. Participants typically respond more slowly and less accurately when the distractor and target arrows point in different directions (incongruent trials) than when they point in the same direction (congruent trials). This *congruency effect* is a measure of overall distractibility.

The congruency effect is often smaller after incongruent trials than after congruent trials. Current views posit that this *congruency sequence effect* (CSE) reflects a mixture of different processes. These include cognitive control processes that minimize distraction from irrelevant stimuli (Botvinick, Braver, Barch, Carter, & Cohen, 2001; Gratton, Coles, & Donchin, 1992), feature integration processes that operate when stimuli and responses repeat in consecutive trials (Hommel, Proctor, & Vu, 2004; Mayr, Awh, & Laurey, 2003), and contingency learning processes that operate when distractors predict target identity with greater-than-chance accuracy (Schmidt & De Houwer, 2011). Consistent with these views, researchers have reported CSEs even in the absence of feature integration and contingency learning confounds in “confound-minimized” protocols (Schmidt & Weissman, 2014; Freitas & Clark, 2015; Hengstler, Holland, van Steenbergen, & van Knippenberg, 2014; Kim & Cho, 2014). These protocols enable researchers to isolate the independent contribution(s) of control processes to the CSE.

Cognitive control accounts of the CSE

There are two competing cognitive control accounts of the CSE. The *attentional shift account* posits that individuals shift their attention toward the target and/or away from the distractor more after incongruent than after congruent trials (Botvinick et al., 2001; Botvinick, 2007; Gratton et al., 1992). In contrast, the *response modulation account* posits that control processes modulate the response signaled by the distractor differently after incongruent trials than after congruent trials (Ridderinkhof, 2002; Stürmer, Leuthold, Soetens, Schröter, & Sommer, 2002; Weissman, Egner, Hawks, & Link, 2015). These accounts are not mutually exclusive, but they sometimes make different predictions that allow them to be distinguished.

Findings from confound-minimized protocols are more consistent with the response modulation account than with the attentional shift account (Weissman, Jiang, & Egner, 2014; Weissman, Egner et al., 2015; Weissman, Colter, Grant, & Bissett, 2017). First, Weissman, Egner et al. (2015) observed a CSE in the absence of an overall congruency effect, which is thought to trigger attentional shifts toward the target (e.g., Botvinick et al., 2001), while employing a 1-second inter-stimulus-interval (ISI) between the distracter and the target[2]. Second, and critically, this CSE was associated with a negative congruency effect (i.e., faster response time in incongruent trials than in congruent trials) after incongruent trials. The authors suggested that, after an incongruent trial, inhibiting the response signaled by the current-trial distractor increases the difficulty of responding to an upcoming target that requires (a) the same, inhibited response (in congruent trials) but not (b) a different, uninhibited response (in incongruent trials). In the absence of an overall congruency effect, selectively increasing the difficulty of responding in congruent trials could lead to a negative congruency effect (Ridderinkhof, 2002; Logan, 1985; Logan & Zbrodoff, 1979; Weissman, Colter, Drake, & Morgan, 2015). In contrast, a negative congruency effect after incongruent trials is inconsistent with the attentional shift account. Even allocating all of one's attention to the target and none to the distracter should lead to the absence of a congruency effect, rather than to a negative congruency effect.

What triggers the CSE?

Two competing variants of the response modulation account posit that different aspects of an incongruent trial trigger the CSE. The *activation-suppression account* posits that the distractor activates an incorrect response, which triggers control processes that inhibit the response signaled by the distractor in the next trial (Ridderinkhof, 2002). This account builds on dual-route architectures wherein distractors and targets activate responses via direct and indirect

---

[2] The appearance of a CSE in the absence of an overall congruency effect suggests that control processes underlying the CSE are not solely dedicated to minimizing distraction. For example, we have suggested that they may index "a mechanism that uses information about a prime, regardless of its task-relevance, to predict an upcoming target response" (Weissman, Colter, Grant, & Bissett, 2017, page 10; Weissman, Colter, et al., 2015).

stimulus-response translation routes, respectively (Kornblum et al., 1990). A CSE occurs when incorrect response activation on the direct route triggers control processes that suppress response activation on this route in the next trial. The *response cueing account* posits that identifying a distractor and a target that cue distinct responses triggers proactive control processes that (a) inhibit the response signaled by the distractor and/or (b) activate the opposite response in the next trial (Alexander & Brown, 2011; Logan, 1985; Logan & Zbrodoff, 1979; Weissman, Colter et al., 2015). Critically, this account posits that identifying two stimuli that cue distinct responses triggers proactive control even in the absence of a dual-route architecture and even when both of the stimuli are relevant targets (Alexander & Brown, 2011).

We recently distinguished between these accounts by using a modified prime-probe arrow task (Weissman, Colter, Grant, & Bissett, 2017, Experiment 1). In each trial, participants viewed either a white or a yellow prime arrow that pointed in the same or opposite direction as an upcoming white probe arrow. A 1000 ms inter-stimulus-interval (ISI) separated the two arrows. In standard trials (66.66%), both arrows were white, and participants indicated the direction in which the probe arrow pointed (left, right, up, or down). In catch trials (33.33%), the prime arrow was yellow and the probe arrow remained white. In these trials, participants indicated (1) the direction in which the prime arrow pointed during the 1000 ms ISI and, then, (2) the direction in which the probe arrow pointed after it appeared. As in standard incongruent trials, the two arrows in catch incongruent trials always cued two distinct responses. However, unlike in standard incongruent trials, the prime arrow in catch incongruent trials could not produce incorrect response activation on the direct route. Indeed, since the researchers instructed participants to respond to the prime arrow in catch trials, the prime arrow was a target and, therefore, could only produce response activation on the indirect (i.e., target-related) route.

The activation-suppression and response cueing accounts make opposing predictions about relative CSE magnitude in (a) standard trials that follow other standard trials and (b) standard trials that follow catch trials. The activation-suppression account predicts a larger CSE in standard trials that follow other standard trials. This is because, as described earlier, only standard incongruent trials activate an incorrect response on the direct route and, hence, trigger control processes that suppress response activation on the direct route in the next trial. In contrast, the response cueing account predicts equivalent, robust CSEs in these sequential trial types. This is because both standard and catch incongruent trials cue two responses. Consistent with the response cueing account, but not with the activation-suppression account, Weissman et al. (2017) observed equivalent, robust CSEs in these two sequential trial types.

However, Weissman et al. (2017) reported a second result that they did not fully explain. Namely, the CSE was numerically larger in sequential catch trials than in sequential standard trials, catch trials followed by standard trials, and standard trials followed by catch trials. Follow-up exploratory analyses revealed that each of these numerical differences was significant[3]. As we describe next, the especially large CSE in sequential catch trials is consistent with two hypotheses.

The unexpected event hypothesis

---

[3] Specifically, the CSE in sequential catch trials (49 ms) was larger than the CSE in (a) sequential standard trials (18 ms; $F(1,31) = 26.967$, $p < 0.001$, $\eta_p^2 = 0.465$), (b) catch trials followed by standard trials (22 ms; $F(1,31) = 11.490$, $p = 0.002$, $\eta_p^2 = 0.270$), and (c) standard trials followed by catch trials (13 ms; $F(1,31) = 18.592$, $p < 0.001$, $\eta_p^2 = 0.375$).

The *unexpected event hypothesis* posits that unexpected events trigger an upregulation of proactive inhibitory control that subsequently inhibits both motor and cognitive processing (Wessel & Aron, 2017). With regard to motor processing, unexpected events, such as novel birdsongs, trigger the same electroencephalography (EEG) signals as action cancellation in the stop-signal task and reduce motor excitability as measured by transcranial magnetic stimulation (TMS) (Wessel & Aron, 2013). With regard to cognitive processing, the degree to which unexpected events trigger the EEG signals above predicts the degree to which unexpected events reduce the quantity of items that can be maintained in working memory (WM) (Wessel & Aron, 2016).

The especially large CSE that Weissman et al. (2017) observed in sequential catch trials is consistent with the unexpected event hypothesis. This is because sequential catch trials were unexpected, rare events that occurred unpredictably in just 11% of the trials. Specifically, perceiving a yellow (rather than a standard white) distractor arrow in the second of two sequential catch trials marked the onset of an especially unexpected, rare event, which may have triggered an increase of proactive inhibitory control. Such an increase could have enhanced ongoing inhibition of the response signaled by the distractor after an initial incongruent catch trial (Weissman, Egner et al., 2015), leading to an especially large CSE. Consistent with this possibility, inhibiting a response triggers control processes underlying the CSE, resulting in a reduced congruency effect in the next trial (Bissett, Grant, & Weissman, 2017; Stürmer et al., 2002).

The episodic retrieval hypothesis

The *episodic retrieval hypothesis* posits the CSE is maximal when the episodic context of the current trial matches the episodic context of the previous trial (Egner, 2014; Hazeltine et al., 2011; Spapé & Hommel, 2008; Spapé & Hommel, 2014). In this view, participants form a memory of each trial, which includes information at different levels of abstraction (Egner, 2014). For example, there is low-level information about the stimulus and response representations that were activated, mid-level information about the relationships between these representations (e.g., congruent or incongruent), and high-level information about the task set (e.g., stimulus-response mappings) and control processes that participants employed. Critically, participants retrieve this memory during the inter-trial-interval (ITI) because the local temporal context (e.g., the previous trial) serves as a cue for guiding expectations about the nature of upcoming events (Egner, 2014). This leads to a larger CSE when participants employ the same task set across two sequential trials than when they do not, even without stimulus or response repetitions (e.g., Ackay & Hazeltine, 2008; Hazeltine et al., 2011).

A number of prior findings are consistent with the episodic retrieval hypothesis. First, the CSE is larger when low-level contextual features (e.g., stimuli, responses, etc.) repeat in sequential trials than when they do not (Hommel et al., 2004; Mayr et al., 2003). Second, when the task structure encourages participants to adopt modality-specific task sets, the CSE is larger when stimuli appear in the same sensory modality in sequential trials than when they appear in different sensory modalities (Hazeltine et al., 2011; Kreutzfeld, Stephan, Willmes, & Koch, 2016). Third, the CSE is larger when participants perform the same task in sequential trials than when they perform different tasks (Kiesel, Kunde, & Hoffmann, 2006).

The especially large CSE that Weissman et al. (2017) observed in sequential catch trials is consistent with the episodic retrieval hypothesis. In sequential catch trials, participants employ the same task set for each stimulus (i.e., “indicate the direction in which the arrow points”) both within and across consecutive trials, which is not possible in the other sequential trial types. For example, in sequential standard trials, participants employ one task set for each prime arrow (i.e., “do *not* indicate the direction which the arrow points”) and another task set for each probe arrow (i.e., “indicate the direction which the arrow points”). Thus, the task set changes *within* each trial, even though participants employ the same task set for (a) the two prime arrows and (b) the two probe arrows *across* consecutive trials. As another example, in standard trials followed by catch trials the task set changes within the first (standard) trial, but not within the second (catch) trial. The task set for the prime arrow also changes *across* these trials because participants employ different task sets for the prime arrow in standard and catch trials. In sum, only in sequential catch trials do participants employ the same task set (i.e., “indicate the direction in which the arrow points”) for all four arrows that appear. The especially large CSE that Weissman et al. (2017) observed in sequential catch trials is, therefore, consistent with the episodic retrieval hypothesis. Notably, confirming this view would indicate that the CSE reflects control processes that do more than minimize distraction. Indeed, there are no distractors in sequential catch trials because each of the four arrows is a target.

The present study

In the present study, we sought to distinguish between the unexpected event and episodic retrieval hypotheses of the especially large CSE Weissman et al. (2017) observed in sequential catch trials. In Experiment 1, we presented standard and catch trials equally often. Thus, sequential catch trials appeared just as frequently as the other three sequential trial types (i.e., in 25% of all trials). We reasoned that this would reduce the degree to which sequential catch trials were unexpected, but not the degree to which participants employed the same task set for all four arrows. In Experiment 2, we investigated the possibility that perceptual differences between standard and catch trials produced our findings in Experiment 1.

Experiment 1

The goal of Experiment 1 was to distinguish between the unexpected event and episodic context hypotheses in relation to the especially large CSE in sequential catch trials. To this end, we modified Experiment 1 of Weissman et al. (2017) such that standard and catch trials appeared equally often (i.e., in 50% of all trials), and, consequently, sequential catch trials appeared as often as the other three sequential trial types (i.e., in 25% of all trials). In this protocol, the unexpected event hypothesis predicts equivalent CSEs in all four sequential trial types because they are equally unexpected. In contrast, the episodic retrieval hypothesis predicts an especially large CSE in sequential catch trials. Indeed, only in these trials does the episodic context/task set remain the same both (a) within each trial and (b) across consecutive trials.

Methods

Participants

Prior to data collection, we conducted power analyses using the partial-eta-squared values associated with the three significant effects in Experiment 1 of Weissman et al. (2017) (see Footnote 1) and an alpha level of 0.05 (G*Power 3.1.9.2; Faul, Erdfelder, Lang, & Buchner, 2007). Our results indicated that 36 participants would provide high power for observing a larger CSE in sequential catch trials than in (a) sequential standard trials (over 99%), (b) standard trials followed by catch trials (over 99%), and (c) catch trials followed by standard trials (93%).

Forty-three undergraduates from the University of Michigan participated for course credit. We excluded data from seven participants for self-reported neurological disorders and/or using psychoactive medications (five participants) and for performing the task with less than 75% accuracy (two participants). None of the remaining 36 participants (17 male, 19 female; 32 right-handed, 4 left-handed; age range: 18-22 years; mean age: 18.86 years, standard deviation of age: 1.05 years) reported any history of head trauma, uncorrected visual or hearing impairments, seizures, or neurological disorders. The University of Michigan's Behavioral Sciences Internal Review Board approved the study.

Experimental Design

The experiment consisted of a 64-trial practice block followed by ten 64-trial test blocks. Each block started and ended with a 2 s fixation cross (0.8° x 0.8°). In each 3 s trial (Fig. 1), there were four sequential events. First, there was a large yellow or white prime arrow that pointed left, right, up, or down (duration, 133 ms; horizontal orientation: 7.8° x 4.7°; vertical orientation, 4.7° x 7.8°). Second, there was a blank screen (duration, 1000 ms). Third, there was a small white probe arrow that pointed left, right, up, or down (duration, 133 ms; horizontal orientation, 2.60°

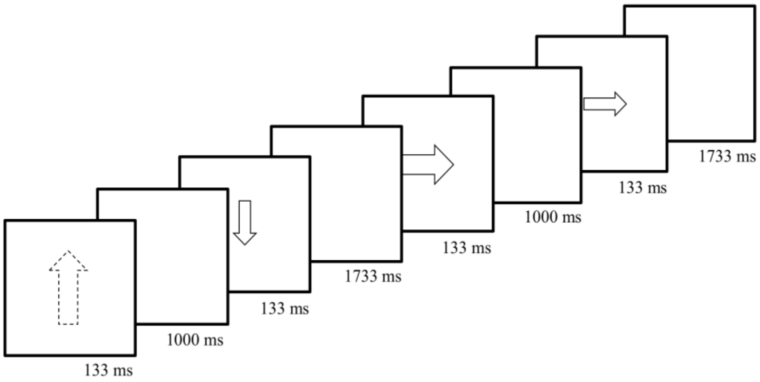


Figure 3-1. Prime-probe task in Experiments 1 and 2

The figure illustrates two trials. In each trial, a large prime arrow preceded a small probe arrow. In half of the trials, the large prime arrow was yellow (as indicated by a dashed line in the figure; e.g., see the arrow on the far left). In the other half, the large prime arrow was white (as indicated by solid lines). The small probe arrow was always white. All stimuli appeared on a black background. The number beneath each box indicates the length of the corresponding trial component in milliseconds (ms).

x 1.6°; vertical orientation, 1.6° x 2.6°). Fourth, there was a second blank screen (duration: 1733 ms). All stimuli appeared on a black background. Participants viewed the stimuli from a distance of approximately 55 cm.

Using the large and small arrow stimuli above, we created sixteen prime-probe pairs. A large *white* prime arrow appeared in eight prime-probe pairs (standard trials) and a large *yellow* prime arrow appeared in eight prime-probe pairs (catch trials). Within each set of eight prime-probe pairs, there were four congruent pairs (left-left, right-right, up-up, and down-down) and four incongruent pairs (left-right, right-left, up-down, and down-up). Thus, in total, eight prime-probe pairs consisted of left-right arrow stimuli (four congruent and four incongruent) and eight prime-probe pairs consisted of up-down arrow stimuli (four congruent and four incongruent). To avoid feature integration confounds (Hommel et al., 2004; Mayr et al., 2003), we presented the left-right arrow stimuli in odd trials and the up-down arrow stimuli in even trials. To avoid contingency learning biases (Schmidt & De Houwer, 2011), we presented each prime-probe pair equally often in every block.

We presented the congruent and incongruent trials in each 64-trial block in a pseudo-randomized order. In particular, we first-order counterbalanced the order of the 16 standard congruent, 16 standard incongruent, 16 catch congruent, and 16 catch incongruent trials. Thus, there were approximately four congruent trials preceded by congruent trials (cC trials), four incongruent trials preceded by congruent trials (cI trials), four congruent trials preceded by incongruent trials (iC trials), and four incongruent trials preceded by incongruent trials (iI trials) when standard trials were followed by standard trials. Analogously, there were approximately four cC, cI, iC, and iI trials when (a) catch trials were followed by catch trials, (b) catch trials were followed by standard trials, and (c) standard trials were followed by catch trials. As no trial preceded the first trial in each block, one of the 16 sequential trial types above only appeared three times. Since we employed a different first-order counterbalanced trial sequence in every block, the underrepresented sequential trial type varied randomly across blocks.

Procedure

We instructed participants to respond immediately to the large prime arrow whenever it was yellow (i.e., during the 1000 ms ISI before the small white probe arrow appeared, see Figure 1) and to the probe arrow in every trial. Participants responded to leftward-pointing arrows by pressing the “F” key (left middle finger), to rightward-pointing arrows by pressing the “G” key (left index finger), to upward-pointing arrows by pressing the “J” key (right middle finger), and to downward-pointing arrows by pressing the “N” key (right index finger) on a QWERTY keyboard. We told participants to respond as quickly and as accurately as possible. Each time that a participant responded incorrectly or failed to respond within 900 ms of arrow onset, the word “Error” appeared centrally for 200 ms. We used the Psychophysics Toolbox to present the stimuli and record participants’ responses (Brainard, 1997).

Data Analyses

We excluded several types of trials. When analyzing mean response time (RT), we excluded practice trials, trials with omitted or incorrect responses, trials following omitted and incorrect responses, and outliers. We defined outliers as trials with RTs at least three standard deviations away from their conditional mean. When analyzing mean error rate (ER), we excluded practice trials, trials following an omitted or incorrect response, and outliers. On average, 7.20% of the trials were errors and 1.02% of the trials were outliers.

We conducted planned contrasts to test our hypothesis that the CSE would be largest when both the previous trial and the current trial were catch trials. In particular, we conducted two-tailed t-tests on mean probe RT to determine whether the CSE was significantly larger in sequential catch trials than in (a) sequential standard trials, (b) catch trials followed by standard trials, and (c) standard trials followed by catch trials.

To reveal the overall pattern of results across all of the conditions, we conducted separate repeated measures analyses of variance (ANOVA) on the mean RT and mean ER data for probe responses. Each ANOVA had four within-participants factors: previous trial type (standard, catch), current trial type (standard, catch), previous trial congruency (congruent, incongruent), and current trial congruency (congruent, incongruent). Table 1 lists the mean RT and mean ER in each experimental condition.

Table 3-1. Mean RT and mean ER in Experiment 1

Experimental condition	cC	cI	iC	iI	Mean	I – C	CSE
Mean RT							
Previous standard, current standard	445 (9)	442 (9)	449 (10)	436 (9)	443 (5)	-8 (5)	11 (5)
Previous catch, current standard	432 (9)	433 (9)	441 (8)	429 (9)	444 (4)	-5 (5)	14 (5)
Previous standard, current catch	419 (9)	428 (9)	428 (9)	422 (9)	424 (4)	1 (4)	15 (6)
Previous catch, current catch	398 (7)	421 (7)	419 (7)	405 (8)	410 (4)	4 (4)	35 (5)
Mean ER							
Previous standard, current standard	4.3 (.8)	3.4 (.6)	3.5 (.6)	3.4 (.7)	3.7 (.3)	-.5 (.6)	-.8 (1)
Previous catch, current standard	8.6 (1.3)	10.1 (1.1)	8.4 (1.2)	9.1 (1.2)	9.1 (.6)	1.1 (.9)	.7 (1.5)
Previous standard, current catch	2.6 (.7)	2.8 (.8)	3.3 (.6)	2.9 (.6)	2.9 (.4)	-.1 (.4)	.6 (.9)
Previous catch, current catch	3.7 (.8)	4.5 (.9)	4.3 (.9)	2.9 (.6)	3.8 (.4)	-.2 (.4)	2.2 (1.5)

Note. cC = Previous trial congruent, current trial congruent; cI = Previous trial congruent, current trial incongruent; iC = Previous trial incongruent, current trial congruent; iI = Previous trial incongruent, current trial incongruent; I-C = Incongruent – Congruent; CSE = Congruency sequence effect, calculated as: (cI - cC) - (iI - iC).

Results

Planned Contrasts

Mean RT

Our main goal was to investigate whether or not the CSE in sequential catch trials is larger than the CSE in the other sequential trial types when sequential catch trials are not rare, unexpected events. We found that the CSE in sequential catch trials (35 ms) was larger than the CSE in (a) sequential standard trials (11 ms; $t(35) = 4.12$, $p < 0.001$), (b) catch trials followed by standard trials (15 ms; $t(35) = 3.18$, $p = 0.003$), and (c) standard trials followed by catch trials

(14 ms; $t(35) = 3.35$, $p = 0.0019$) (Fig. 2). These findings appear more consistent with the episodic retrieval hypothesis than with the unexpected event hypothesis.

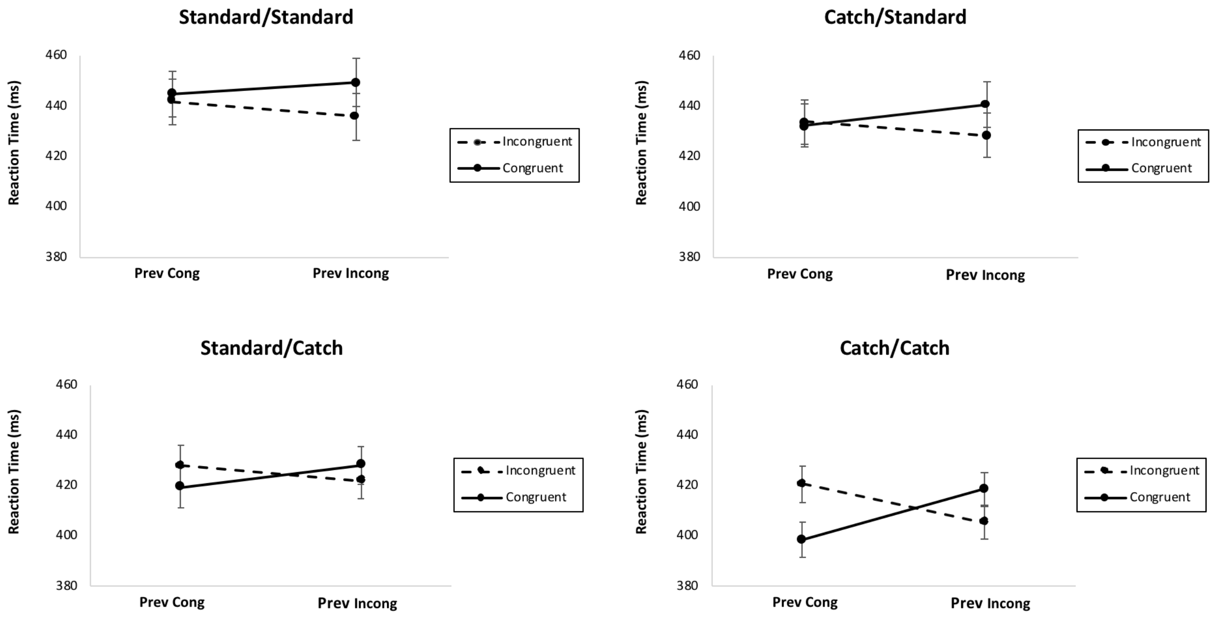


Figure 3-2. The CSE in the four trial sequences of Experiment 1
Previous trial congruency appears on the x-axis (Prev Cong: previous congruent trial; Prev Incong: previous incongruent trial). Current trial congruency is distinguished by line type (Incongruent: solid line; Congruent: dashed line). Reaction time (in ms) appears on the y-axis.

Mean ER

Overall, we did not observe a significant CSE in the mean ER data (see below). Thus, we did not conduct planned contrasts analogous to those we conducted on the mean RT data.

Repeated-measures ANOVAs

To reveal the overall pattern of results, we now present the results of the repeated-measures ANOVAs.

Mean RT

There were two significant main effects. First, there was a main effect of previous trial type, $F(1,35) = 66.513$, $p < 0.001$, $\eta_p^2 = 0.655$, because mean RT was longer after standard trials (434 ms) than after catch trials (422 ms). Second, there was a main effect of current trial type, $F(1,35) = 22.111$, $p < 0.001$, $\eta_p^2 = 0.387$, because mean RT was longer in standard trials (438 ms) than in catch trials (418 ms).

There were also three significant two-way interactions. First, there was an interaction between previous trial type and current trial type, $F(1,35) = 4.159$, $p = 0.049$, $\eta_p^2 = 0.106$: mean RT was longer in standard trials than in catch trials, but this difference was smaller after standard trials (19 ms) than after catch trials (23 ms). Second, there was an interaction between previous trial congruency and current trial congruency, $F(1,35) = 54.500$, $p < 0.001$, $\eta_p^2 = 0.609$, because the congruency effect was smaller after incongruent trials (-11 ms; $F(1,35) = 7.321$, $p < 0.05$, $\eta_p^2 = 0.173$) than after congruent trials (7 ms; $F(1,35) = 4.042$, $p > 0.05$, $\eta_p^2 = 0.104$). Third, there was an interaction between current trial type and current trial congruency, $F(1,35) = 6.087$, $p = 0.019$, $\eta_p^2 = 0.148$, because the congruency effect was smaller in standard trials (-7 ms) than in catch trials (3 ms).

Finally, there were a pair of significant three-way interactions. First, there was an interaction among current trial type, previous trial congruency, and current trial congruency, $F(1,35) = 6.291$, $p = 0.017$, $\eta_p^2 = 0.152$: the CSE was larger in catch trials (25 ms; $F(1,35) = 40.878$, $p < 0.001$, $\eta_p^2 = 0.539$) than in standard trials (12 ms; $F(1,35) = 13.667$, $p < 0.001$, $\eta_p^2 = 0.281$). Second, there was an interaction among previous trial type, previous trial congruency, and current trial congruency, $F(1,35) = 7.091$, $p = 0.012$, $\eta_p^2 = 0.168$: the CSE was larger after catch trials (25 ms; $F(1,35) = 56.018$, $p < 0.001$, $\eta_p^2 = 0.615$) than after standard trials (13 ms; $F(1,35) = 13.513$, $p < 0.001$, $\eta_p^2 = 0.279$). In short, the CSE was larger when either the previous trial or the current trial was a catch trial.

Interestingly, we did not observe a significant four-way interaction among previous trial type, current trial type, previous trial congruency, and current trial congruency, $F(1,35) = 3.404$, $p = 0.073$, $\eta_p^2 = 0.089$. However, there was a non-significant trend toward such an effect. No other effects were significant.

Mean ER

There were two significant main effects. First, there was a main effect of previous trial type, $F(1,35) = 72.685$, $p < 0.001$, $\eta_p^2 = 0.675$, because mean ER was lower after standard trials (3.3%) than after catch trials (6.5%). This main effect, which is opposite to the analogous one observed in mean RT, suggests that participants traded speed for accuracy after standard trials. Second, there was a main effect of current trial type, $F(1,35) = 15.638$, $p < 0.001$, $\eta_p^2 = 0.309$, because mean ER was greater in standard trials (6.4%) than in catch trials (3.4%).

There was also a significant two-way interaction. In particular, there was an interaction between previous trial type and current trial type, $F(1,35) = 31.568$, $p < 0.001$, $\eta_p^2 = 0.474$: mean ER was greater in standard trials than in catch trials, but this difference was smaller after standard trials (0.8%) than after catch trials (5.6%).

Finally, as in the mean RT data, we did not observe a significant four-way interaction among previous trial type, current trial type, previous trial congruency, and current trial congruency, $F(1,35) < 1$. No other effects were significant.

Discussion

In the mean RT data, the CSE was larger in sequential catch trials than in the other three sequential trial types even when these trial types appeared equally often. This result is consistent with the episodic retrieval hypothesis, which posits the CSE is largest when participants employ the same task set both within each trial and across consecutive trials. In contrast, this result is inconsistent with the unexpected event hypothesis, which posits the CSE is largest when a particular trial type is unexpected or rare. Thus, our findings appear more consistent with the episodic retrieval hypothesis than with the unexpected event hypothesis.

However, presenting the prime and probe arrows in different colors (i.e., yellow and white) in each of two sequential catch trials may have contributed to our findings. For instance, CSE magnitude increases when a distractor and a target activate distinct perceptual representations, which may trigger heightened control (Verbruggen, Notebaert, Liefooghe, & Vandierendonck, 2006). Thus, presenting the prime and probe arrows in different colors, which activate distinct color representations, may have increased the CSE.

CSE magnitude also increases when the display includes colors that trigger relatively high (versus low) levels of alertness. Specifically, Wang, Zhao, Xue, & Chen (2016) reported a larger CSE when they presented task stimuli on red backgrounds relative to blue backgrounds, and that this effect reflected increased alertness in trials with red backgrounds. Notably, the authors based their prediction that red backgrounds would increase alertness more than blue backgrounds on the results of a prior study wherein participants rated the perceived hazard of signal words (e.g., DANGER) more highly when those words appeared in red as compared to blue (Braun & Silver, 1995). Wang et al. (2016) interpreted these findings as suggesting that red increases alertness more than blue, and their study confirmed this hypothesis. Critically, participants perceive yellow as conveying greater hazard than white (Dunlap et al., 1986). Thus, prime arrows presented in yellow may increase alertness more than prime arrows presented in white, and this difference may yield a larger CSE in sequential catch trials (wherein two yellow primes appear) than in the other sequential trial types (wherein one or zero yellow primes appear).

Experiment 2

The goal of Experiment 2 was to rule out the color confound described above. To accomplish this objective, we asked a new group of 32 participants to complete the same task as in Experiment 1 with one exception. Specifically, we instructed participants to respond to both the prime arrow and the probe arrow in each trial, regardless of whether the prime arrow was white or yellow. Under these conditions, the episodic retrieval hypothesis predicts large, equivalent CSEs in the four sequential trial types. Specifically, the requirement to respond to both arrows in every trial should lead participants to employ the same task set for all four arrows in every possible pair of consecutive trials, thereby maximizing the CSE in every sequential trial type. In contrast, if the color confound triggered the especially large CSE that we observed in sequential catch trials, then we should still observe this effect in Experiment 2. Indeed, the prime and probe arrows still appear in different colors (i.e., yellow and white, respectively) more often in sequential catch trials than in any of the other sequential trial types.

Methods

Participants

To be consistent with Experiment 1, we collected usable data from 36 new participants[4]. Thirty-six undergraduates at the University of Michigan participated for course credit. None of the participants reported any history of head trauma, uncorrected vision or hearing impairments, seizures, or neurological disorders (15 male, 21 female; 33 right-handed, 3 left-handed; age range, 18-24 years; mean age, 19 years; standard deviation of age, 1.41 years). The University of Michigan's Behavioral Sciences Internal Review Board approved the study.

Experimental Design

The experimental design was identical to that in Experiment 1.

Procedure

The procedure was identical to that in Experiment 1 with one exception: we instructed participants to respond to both the prime arrow and the probe arrow in every trial.

Data Analyses

The data analyses were identical to those in Experiment 1. On average, 8.80% of the trials were errors and 0.78% of the trials were outliers. Table 2 lists the mean RT and mean ER in each experimental condition.

---

[4] Power analyses based on the results of Experiment 1 (alpha = 0.05) also indicated that 36 participants provided high power for observing a larger CSE in sequential catch trials than in (a) sequential standard trials (over 99%), (b) catch trials followed by standard trials (90%), and (c) standard trials followed by catch trials (93%).

Table 3-2. Mean RT and mean ER in Experiment 2

Experimental condition	cC	cI	iC	iI	Mean	I – C	CSE
Mean RT							
Previous standard, current standard	384 (7)	405 (9)	412 (9)	379 (8)	395 (8)	-6 (4)	54 (5)
Previous catch, current standard	385 (8)	404 (9)	409 (8)	382 (8)	395 (8)	-4 (4)	46 (5)
Previous standard, current catch	388 (7)	406 (10)	410 (8)	383 (7)	397 (8)	-5 (5)	45 (7)
Previous catch, current catch	387 (7)	403 (9)	411 (8)	385 (8)	396 (8)	-5 (5)	42 (6)
Mean ER							
Previous standard, current standard	5.9 (.9)	9.3 (1.5)	8.7 (1.3)	4.4 (.8)	7.1 (.6)	-.5 (.8)	7.7 (1.9)
Previous catch, current standard	5.5 (1.0)	7.8 (1.5)	7.3 (1.1)	5.0 (1.0)	6.4 (.6)	0.0 (1.0)	4.5 (1.6)
Previous standard, current catch	6.2 (1.1)	6.4 (1.2)	10 (1.4)	5.4 (1.0)	7.0 (.6)	-2.2 (.7)	4.8 (1.5)
Previous catch, current catch	6.9 (1.0)	6.6 (1.4)	10.6 (1.7)	5.0 (.9)	7.3 (.7)	-3.0 (1.0)	5.3 (1.9)

Note. cC = Previous trial congruent, current trial congruent; cI = Previous trial congruent, current trial incongruent; iC = Previous trial incongruent, current trial congruent; iI = Previous trial incongruent, current trial incongruent; I-C = Incongruent – Congruent; CSE = Congruency sequence effect, calculated as: (cI - cC) - (iI - iC).

Results

Planned Contrasts

Mean RT

Planned contrasts revealed that the CSE in sequential catch trials was not significantly larger than the CSE in any of the other three sequential trial types (Fig. 3). First, and somewhat surprisingly, the CSE in sequential catch trials (42 ms) was significantly smaller than the CSE in sequential standard trials (54 ms; $t(35) = -2.14$, $p = 0.039$). Second, the CSE in sequential catch trials did not differ from the CSE in (a) catch trials followed by standard trials (45 ms; $t(35) = -$

0.64, $p > 0.05$), or (b) standard trials followed by catch trials (46 ms; $t(35) = -0.78$, $p > 0.05$). Although the larger CSE in sequential standard trials than in sequential catch trials was unexpected, on balance these findings appear more consistent with the episodic retrieval hypothesis than with an influence of color confounds on the CSE.

Mean ER

Unlike in Experiment 1, we observed a significant CSE in the mean ER data. Thus, we conducted planned contrasts as in the mean RT data. The CSE in sequential catch trials (5.3%) did not differ from the CSE in sequential standard trials (7.7%; $t(35) = -1.090$, $p > 0.05$), catch trials followed by standard trials (4.8%; $t(35) = 0.308$, $p > 0.05$), or standard trials followed by catch trials (4.5%; $t(35) = 0.229$, $p > 0.05$).

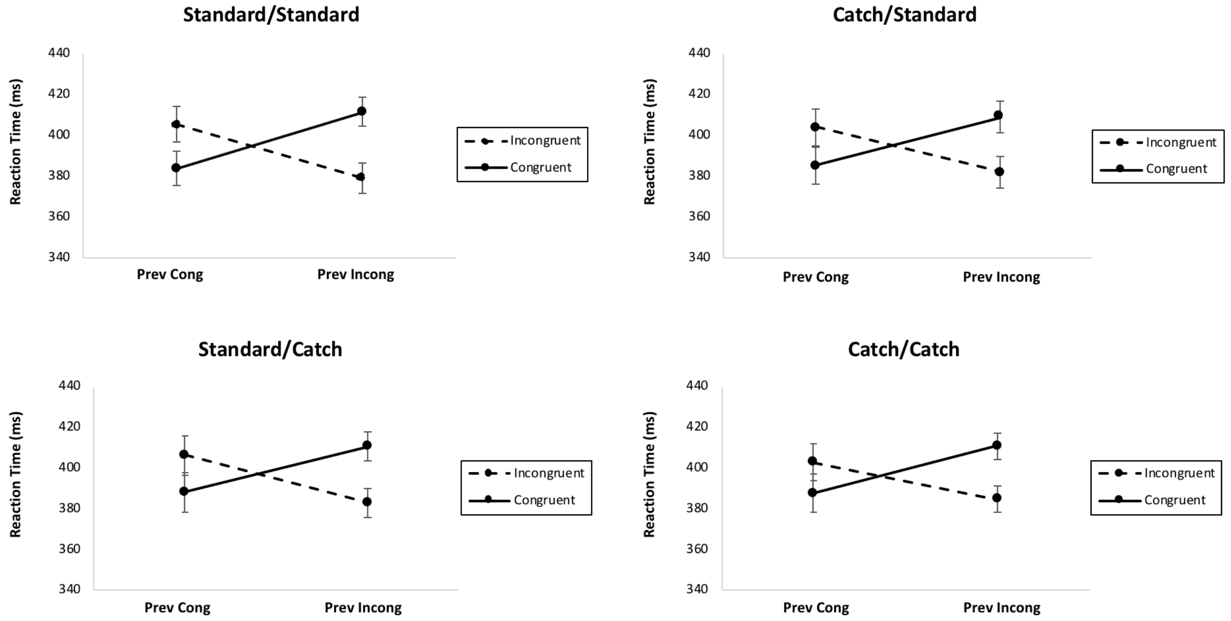


Figure 3-3. The CSE in the four trial sequences of Experiment 2
Trial congruency appears on the x-axis (Prev Cong: previous congruent trial; Prev Incong: previous incongruent trial). Current trial congruency is distinguished by line type (Incongruent: solid line; Congruent: dashed line). Reaction time (in ms) appears on the y-axis.

Repeated-measures ANOVAs

Mean RT

There was one significant effect. Specifically, there was a significant two-way interaction between previous trial congruency and current trial congruency, $F(1,35) = 108.305$, $p < 0.001$, $\eta_p^2 = 0.756$: the congruency effect was smaller after incongruent trials (-28 ms; $F(1,35) = 48.260$, $p < 0.001$, $\eta_p^2 = 0.580$) than after congruent trials (18 ms; $F(1,35) = 15.100$, $p < 0.001$, $\eta_p^2 = 0.301$).

Finally, consistent with the results of the planned contrasts, we did not observe a significant four-way interaction among previous trial type, current trial type, previous trial congruency, and current trial congruency, $F(1,35) < 1$. No other effects were significant.

Mean ER

There was one significant main effect. Specifically, there was a main effect of current trial congruency, $F(1,35) = 7.025$, $p = 0.012$, $\eta_p^2 = 0.167$: mean ER was smaller in incongruent trials (6.2%) than in congruent trials (7.6%) (i.e., there was a reverse congruency effect).

There were also three significant two-way interactions. First, there was an interaction between current trial type and current congruency, $F(1,35) = 8.902$, $p = 0.005$, $\eta_p^2 = 0.203$, because the congruency effect was smaller in catch trials (-2.6%) than in standard trials (-0.2%). Second, there was an interaction between current trial type and previous congruency, $F(1,35) = 4.834$, $p = 0.035$, $\eta_p^2 = 0.121$: the degree to which mean ER was higher in catch trials than in standard trials (6.7%) was reduced when the previous trial was congruent (-0.6%) as compared to incongruent (1.4%). Third, there was an interaction between previous trial congruency and current trial congruency, $F(1,35) = 23.493$, $p < 0.001$, $\eta_p^2 = 0.402$, because the congruency effect was smaller after incongruent trials (-4.2%; $F(1,35) = 34.41$, $p < 0.001$, $\eta_p^2 = 0.496$) than after congruent trials (1.4%; $F(1,35) = 2.791$, $p > 0.05$, $\eta_p^2 = 0.074$).

Finally, consistent with the results of the planned contrasts we conducted earlier on the mean ER data, we did not observe a significant four-way interaction among previous trial type, current trial type, previous trial congruency, and current trial congruency, $F(1,35) = 1.171$, $p = 0.287$, $\eta_p^2 = 0.032$. No other effects were significant.

Across-experiment analyses

Consistent with the episodic retrieval hypothesis, our findings thus far suggest that maintaining the same task set both within and across sequential catch trials engenders an especially large CSE. We therefore conducted across-experiment analyses to test a further prediction of this hypothesis. Specifically, this hypothesis predicts an interaction between Experiment (1, 2) and Condition (sequential catch trials, average of the three other sequential trial types). More concretely, the CSE in sequential catch trials should be relatively large in both experiments, because participants respond to both arrows in each trial. In contrast, the CSE in the other three sequential trial types should be larger in Experiment 2 (wherein participants respond to both arrows in each trial) than in Experiment 1 (wherein participants do not respond to the prime arrow in one or both trials).

To test this prediction, we conducted a mixed ANOVA on mean probe RT in which Experiment (1, 2) was a between-participants factor, Condition (sequential catch trials, average of the other three sequential trial types) was a within-participants factor, and the CSE was the dependent measure. There was a main effect of Condition, $F(1,70) = 5.003$, $p = 0.028$, $\eta_p^2 = 0.067$, because the CSE was larger in sequential catch trials (46 ms) than (on average) in the other three sequential trial types (36 ms). Further, consistent with the episodic retrieval hypothesis, we observed a significant two-way interaction between Experiment and Condition, $F(1,70) = 18.193$, $p < 0.001$, $\eta_p^2 = 0.206$. As predicted, the CSE in sequential catch trials did not vary across the two experiments (35 ms vs 42 ms; $F(1,70) = 0.727$, $p > 0.05$, $\eta_p^2 = 0.010$). In contrast, the average CSE in the other three sequential trial types was larger in Experiment 2 than in Experiment 1 (45 ms vs 24 ms; $F(1,70) =14.82$, $p < 0.001$, $\eta_p^2 = 0.175$).

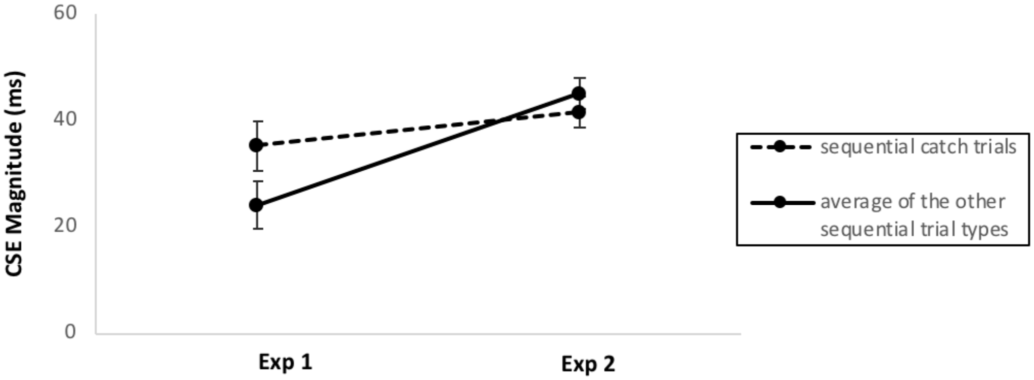


Figure 3-4. The CSE in Experiments 1 and 2
The solid line indicates the CSE in sequential catch trials. The dashed line indicates the CSE averaged across the other three sequential trial types (see text for details). Experiment (1 or 2) is indicated on the x-axis. CSE magnitude (in ms) is indicated on the y-axis.

Discussion

Unlike in Experiment 1, the CSE in sequential catch trials was no larger than the CSEs in the other sequential trial types. Further, across-experiment analyses revealed that while the CSE in sequential catch trials did not vary across Experiments 1 and 2, the average CSE in the other three sequential trial types was larger in Experiment 2 than in Experiment 1. These findings are consistent with the episodic retrieval hypothesis, which posits that CSE magnitude varies with the degree to which participants employ the same task set both within and across trials. In contrast, these findings weigh against the color confound hypothesis, which always predicts the largest CSE in sequential catch trials. Finally, these findings provide further evidence against the unexpected event hypothesis by showing that the CSE in sequential catch trials remains robust even when these trials are frequent, expected events that occur 100% of the time.

General Discussion

The present findings indicate that turning the distractor (i.e., the prime) in the prime-probe task into a target increases the CSE, even when doing so does not constitute a rare, unexpected event. This outcome is more consistent with the episodic retrieval account of the especially large CSE that Weissman et al. (2017) observed in sequential catch trials than with the unexpected event account. More broadly, it extends the view that task sets determine the boundaries of control processes underlying the CSE by revealing that this phenomenon appears even in the absence of irrelevant stimuli. The emergence of the CSE under such conditions reveals that the control processes underlying this phenomenon contribute to cognition in ways that extend beyond minimizing distraction from irrelevant stimuli. As we describe next, these findings have important implications for current views of the CSE.

The present findings are consistent with the episodic retrieval hypothesis

The episodic retrieval hypothesis predicts the largest CSE when participants maintain the same episodic context/task set while identifying sequential stimuli (Egner, 2014; Fischer, Plessow, Kunde, & Kiesel, 2010*;* Hazeltine et al., 2011; Spapé & Hommel, 2008; Spapé & Hommel, 2014; Weissman, Hawks, & Egner., 2016). In Experiment 1, participants could employ the same task set (i.e., “indicate the direction in which the arrow points”) while identifying the four arrows in sequential catch trials but needed to employ different task sets in the other three sequential trial types. In Experiment 2, participants could employ the same task set when identifying the four arrows in all sequential trial types. Consistent with the episodic retrieval hypothesis, we observed an especially large CSE only in sequential catch trials in Experiment 1 but in all sequential trial types in Experiment 2.

These findings complement prior data indicating that maintaining the same episodic context across trials amplifies the CSE. For example, the CSE is larger when two sequential trials involve the same stimuli (e.g., Mayr et al., 2003; Spapé & Hommel, 2008; Weisman et al., 2016), sensory modality (e.g., Yang, Nan, Zheng, Wu, Li, & Liu, 2017), and/or task set (e.g., Ackay & Hazeltine, 2008; Hazeltine et al., 2011; Kiesel, Kunde, & Hoffmann, 2006) than when they do not. The present findings extend these prior results by showing for the first time that maintaining the same episodic context/task set *within* each trial also increases the size of the CSE, and that this effect occurs even when the prime is a task-relevant target.

As such, the present findings fit with our prior suggestion that the CSE "may index a mechanism that uses information about a prime, regardless of its task-relevance, to predict an upcoming target response" (Weissman et al., 2017, page 10; see also, Weissman, Colter et al., 2015). The view that control processes underlying the CSE incorporate information about distractors (i.e., previous-trial congruency) to predict whether upcoming targets will require congruent or incongruent responses is not new (Duthoo, Wühr, & Notebaert, 2013; Jiang, Heller, & Egner, 2014; Logan & Zbrodoff, 1979; Hommel & Wiers, 2017; Stürmer et al., 2002). To our knowledge, however, prior researchers have not suggested that control processes engender CSEs even when a trial involves two targets, rather than a distractor and a target.

One possible explanation is that control processes underlying the CSE are similar to control processes that develop expectations about upcoming response repetitions and response alternations in 2-AFC tasks (e.g., Huettel, Mack, & McCarthy, 2002). In each trial of such tasks, one of two possible stimuli requiring a distinct response randomly appears. Critically, control processes appear to predict (1) a response repetition after trials in which the stimulus repeats or (2) a response alternation after trials in which the stimulus alternates. Indeed, participants respond more slowly when new stimuli violate these expectations than when they do not.

Similarly, control processes underlying the CSE may predict whether the prime and the probe in each trial will require the same response (a response repetition) or different responses (a response alternation), based on whether similar stimuli in the previous trial required the same or different responses. Such predictions typically cannot exceed greater-than-chance accuracy when trial congruency varies randomly in laboratory studies. However, they may still occur because they are often more accurate in the real world. Indeed, memories of how we attended and responded to sequential stimuli in the recent past (e.g., smiling politely when each of two colleagues said hello) often accurately predict how we should respond to similar sequential stimuli in the near future (Egner, 2014). Future work investigating the potential link between control processes that (1) underlie the CSE and (2) develop expectations for response repetitions and response alternations in random sequences may therefore be a fruitful avenue for future research (see Duthoo et al., 2013 and Jimenez & Mendez, 2013, 2014 for related work).

The present findings are inconsistent with the unexpected event hypothesis

The unexpected event hypothesis predicts the largest CSE in sequential catch trials when participants expect these trials to a lesser degree than the other sequential trial types, as in Weissman et al. (2017). In Experiment 1, however, we observed a larger CSE in sequential catch trials than in each of the other three sequential trial types even when sequential catch trials appeared with the same unpredictable timing and frequency as the other sequential trial types and, hence, were unexpected to the same degree. This result suggests that the especially large CSE that Weissman et al. (2017) observed in sequential catch trials is not specific to task protocols wherein sequential catch trials are unexpected events.

It is important to mention, however, that this result does not invalidate the unexpected event hypothesis more generally for at least two reasons. First, Wessel & Aron (2017) suggest that unexpected events may not recruit proactive inhibitory control when they require strategic adaptations to ongoing behavior, such as in task-switching protocols. Along these lines, responding to the unexpected prime arrow in catch trials requires participants to abandon the typical task set for the prime arrow (i.e., "do *not* indicate the direction in which the prime arrow points) in favor an atypical task set (i.e., do indicate the direction in which the prime arrow points). Thus, catch trials, though unexpected, may not trigger proactive inhibitory control. Second, unexpected events may trigger proactive inhibitory control processes that do not interact with the control processes underlying the CSE. This possibility is difficult to reconcile with data suggesting the CSE indexes control processes that modulate (e.g., inhibit) the response signaled by the distractor (e.g., Stürmer et al., 2002; Weissman, Egner et al., 2015). However, it could potentially explain why sequential catch trials are associated with an especially large CSE regardless of whether they are unexpected relative to the other sequential trial types. Most important for present purposes, our findings favor the episodic retrieval hypothesis over the unexpected event hypothesis, because they reveal an especially large CSE in sequential catch trials even when these trials are not less expected than the other sequential trial types.

The present findings are not driven by color confounds

In Experiment 1, the especially large CSE that we observed in sequential catch trials could have indexed the fact that the prime and probe appeared in different colors (i.e., yellow and white, respectively) in two consecutive trials. Contrary to this possibility, in Experiment 2 the CSE in sequential catch trials was not larger than the CSEs in the other sequential trial types. This result argues against the possibility that presenting the prime and probe in different colors triggered control processes that increase the CSE (Verbruggen et al., 2006). It also argues against the possibility that presenting a yellow prime arrow twice in sequential catch trials, relative to once or not at all in the other sequential trial types, heightened alertness, which increases the CSE (Wang et al., 2016).

Nonetheless, the results of Experiment 2 suggest that the color in which the prime and probe appear influences CSE magnitude to some degree. In particular, we observed a larger CSE in sequential standard trials (wherein all four arrows appeared in white) than in sequential catch trials (wherein the primes appeared in white and the probes appeared in yellow). Future studies could investigate the nature of this effect. For example, it may index interactions between feature (e.g., color) repetitions and control processes that increase the CSE (Weissman et al., 2016).

Broader implications

The present findings add to a growing body of evidence indicating that the conflict monitoring account does not provide an adequate explanation of the CSE (Weissman et al., 2014; Weissman, Egner et al., 2015; Weissman et al., 2017). According to this account, the presence of response conflict in an incongruent trial triggers control processes that increase attention to task-relevant stimuli and reduce attention to task-irrelevant stimuli, thereby reducing the congruency effect in the next trial (Botvinick et al., 2001). Contrary to this account, however, the present findings revealed CSEs in the absence of an overall congruency effect, which is the behavioral signature of response conflict (Yeung, Cohen, & Botvinick, 2011). Further, the CSEs that we observed were associated with a negative congruency effect (i.e., faster responses in incongruent trials than in congruent trials) after incongruent trials. A negative congruency effect is inconsistent with a reduction of attention to the distractor, which, at best, should lead to the absence of a congruency effect. These findings further reveal that the conflict monitoring account does not provide a complete explanation of the confound-minimized CSE.

Our findings are also interesting to consider from the view that the CSE indexes a passive carryover of attentional control settings from one trial to the next (Hubbard, Kuhns, Schäfer, & Mayr, 2017). According to this view, the control settings underlying performance in one trial carry over to the next trial in the absence of an active conflict regulation mechanism. Therefore, experiencing high conflict in one trial should be a harbinger of high conflict (due to passive carryover) rather than low conflict (due to active regulation) in the next trial. Consistent with this hypothesis, Hubbard et al. (2017) reported that participants who experienced high conflict from a sudden onset distractor in an eye movement task, as indexed by fixating the distractor, tended to fixate (rather than avoid fixating) the onset distractor in the next trial.

We did not design the present study to rigorously test whether a passive carryover of control settings underlies the CSE in the prime-probe task. Nonetheless, prior data suggest that the CSE in this task indexes a modulation of the response signaled by the prime (e.g., Weissman, Egner et al., 2015). The CSE in this task could therefore reflect a passive carryover of control settings that (a) enhance the response signaled by the prime in congruent trials or (b) inhibit the response signaled by the prime in incongruent trials. This view predicts faster responses to primes after congruent trials than after incongruent trials.

Contrary to this view, an exploratory analysis of the data from Experiment 2 revealed *slower* (and more accurate) responses to primes after congruent trials than after incongruent trials[5]. We speculate that control processes raised the threshold for responding to the prime after participants executed the same response twice (once to the prime, once to the probe) in a previous congruent trial to prevent the system from executing the same response a third time when the upcoming prime appeared (which was never a correct response in our trial sequences). Then, *after* participants responded to the prime, we speculate that control processes enhanced or inhibited the response signaled by the prime – depending on whether the previous trial was congruent or incongruent – to prepare a predicted response to the upcoming probe. Whether or not this speculation is accurate, our exploratory analyses suggest that a passive carryover of attentional control settings may not provide a complete account of CSEs. Thus, researchers should conduct additional studies to evaluate the passive carryover view in more detail.

The present results are also relevant to our understanding of “real world” congruency sequence effects. For example, the negative influence of a head fake on performance is reduced after a previous head fake in basketball players, but not in soccer players or non-athletes who lack basketball expertise (Weigelt, Güldenpenning, Steggemann-Weinrich, Alaboud, & Kunde, 2017). The researchers interpreted these findings as indicating that basketball expertise allows a

---

[5] Both of these effects were significant. First, participants responded more slowly to primes after congruent trials than after incongruent trials (426 ms vs. 412 ms; $F(1, 35) = 83.527$, $p < 0.001$, $\eta_p^2 = 0.705$). Second, participants made fewer errors when responding to primes after congruent trials than when responding to primes after incongruent trials (9.6% vs. 12.7%; $F(1, 35) = 19.966$, $p < 0.001$, $\eta_p^2 = 0.363$). Thus, participants appear to have traded speed for accuracy when responding to primes after congruent trials.

player to suppress an opponent's head orientation and/or to increase attention to the relevant basketball. The present findings from the prime-probe task, which shares important features with a head fake as described in the Introduction, are consistent with this interpretation. They also suggest the possibility that, when experiencing the second of two consecutive head fakes, a basketball player modulates response activation engendered by the second head fake in ways that facilitate responding to a basketball pass in the opposite direction. Given that minimizing the impact of various types of fakes is important in professional sports, future studies investigating the potential contributions of control processes underlying the CSE to this ability could have important real-world implications (e.g., for training athletes).

Limitations

One may wonder whether the present findings reflect control processes that contribute to performance only when a long ISI separates a prime from a probe. Contrary to this possibility, the CSE in the prime-probe task does not vary with whether a short (33 ms) or a long (1000 ms) ISI separates the prime and the probe (Weissman, Egner, et al., 2015). Thus, the CSEs we have observed are unlikely to index control processes whose recruitment is specific to long ISIs.

One may also wonder whether the especially large CSEs that we have observed reflect a shift of attention toward the target or a modulation of the response signaled by the prime after incongruent trials. As described earlier, prior findings from confound-minimized protocols (e.g., negative congruency effects after incongruent trials) are more consistent with the response modulation account than with the attentional shift account (Weissman, Egner, et al., 2015). It remains unclear, however, whether, after incongruent trials, control process inhibit the response signaled by the prime, activate the opposite response, or do a little of both. By employing methodologies that provide direct measures response activation, future researchers could aim to distinguish among these alternatives.

Finally, one may wonder whether responding to the prime in catch incongruent trials engendered greater incorrect response activation (i.e., response conflict) than not responding to the prime in standard incongruent trials, thereby triggering an especially large CSE in sequential catch trials. In line with this possibility, the congruency effect in Experiment 1 was larger in catch trials (3 ms) than in standard trials (-7 ms). However, the size of the confound-minimized CSE does not vary with the size of the congruency effect (e.g., Weissman, Egner, et al., 2015; Weissman, Hawks, & Egner, 2016). As a rather extreme example, Weissman et al. (2015, Experiment 3) reported that CSE magnitude does not vary with whether there is a very large congruency effect or no congruency effect (i.e., 78 ms versus 0 ms). Similarly, in the present study we observed robust CSEs in in the complete absence of overall congruency effects. Thus, the slightly larger congruency effect in catch trials than in standard trials that we observed is unlikely to explain the especially large CSE that we observed in sequential catch trials.

Conclusion

The present findings indicate that the especially large CSE Weissman et al. (2017) observed in sequential catch trials is more consistent with the episodic retrieval hypothesis than with the unexpected event hypothesis. Furthermore, to our knowledge, they are the first to reveal that maintaining the same task set *within* each trial (i.e., by turning distractors into targets) increases CSE magnitude. Future studies that further investigate the influence of task sets on the CSE may shed additional light on this important measure of adaptive cognitive control.

Chapter 4 Task Sets Serve as Boundaries for the Congruency Sequence Effect

Abstract

Cognitive control processes that enable purposeful behavior are often context-specific. A teenager, for example, may inhibit the tendency to daydream at work but not in the classroom. However, the nature of contextual boundaries for cognitive control processes remains unclear. We therefore revisited an ongoing controversy over whether such boundaries reflect (1) an attentional reset that occurs whenever a context-defining (e.g., sensory) feature changes or (2) a disruption of episodic memory retrieval that occurs only when the updated context-defining feature is linked to a different task set. To distinguish between these hypotheses, we employed a cross-modal distractor-interference task to determine precisely when changing a salient context-defining feature – the sensory modality in which task stimuli appear – bounds control processes underlying the congruency sequence effect (CSE). Consistent with the task set hypothesis, but not with the attentional reset hypothesis, Experiments 1 and 2 revealed that changing the sensory modality in which task stimuli appear eliminates the CSE only when the task structure enables participants to form modality-specific task sets. Experiment 3 further revealed that such "modality-specific" CSEs are associated with orienting attention to the sensory modality in which task stimuli appear, which may facilitate the formation of a modality-specific task set. These findings support the view that task sets serve as boundaries for the CSE.

Introduction

Control processes that enable purposeful behavior are often specific to the situational context in which they are recruited. For example, after being distracted by someone talking during a movie, a moviegoer may actively prepare to ignore that person's voice. This type of preparation, however, typically ends once the movie is over and the house lights come on. Such adaptive, context-dependent behavior is an integral aspect of *cognitive control*, which allows individuals to successfully engage in goal-oriented behavior (Miller & Cohen, 2001).

Researchers often use distractor-interference tasks to investigate cognitive control (Eriksen & Eriksen, 1974; Eriksen & Schultz, 1979; Simon & Rudell, 1967; Stroop, 1935). For example, in the prime-probe task, researchers instruct participants to respond to a target while ignoring a preceding distractor. In each trial, the distractor signals the same response as the upcoming target (congruent trials) or a different response (incongruent trials). In general, participants perform more slowly and less accurately in incongruent relative to congruent trials. This phenomenon is known as the congruency effect.

Notably, the congruency effect is smaller after incongruent trials than after congruent trials (Gratton, Coles, & Donchin, 1992). Over the past twenty-five years, there has been an extensive debate over what this *congruency sequence effect* (CSE) reflects. Some researchers posit that it reflects control processes that minimize distraction after incongruent relative to congruent trials (Botvinick, Nystrom, Fissell, Carter, & Cohen, 1999; Gratton et al., 1992; Ullsperger, Bylsma, & Botvinick, 2005). In contrast, others argue that it reflects feature integration (e.g., stimulus repetition) and/or contingency learning (e.g., stimulus frequency) processes that are often confounded with trial congruency (Hommel, Proctor, & Vu, 2004; Mayr, Awh, & Laurey, 2003; Schmidt & De Houwer, 2011). Recently, however, researchers have shown that all of these processes contribute to the CSE (Braem, Abrahamse, Duthoo, & Notebaert, 2014; Mordkoff, 2012; Schmidt, 2013; Schmidt & Weissman, 2014; Weissman, Hawks, & Egner, 2016). Moreover, researchers have developed protocols that yield robust "control-driven" CSEs in the absence of these confounds (Schmidt & Weissman, 2014; Weissman, Egner, Hawks, & Link, 2015; Weissman, Jiang, & Egner, 2014). Thus, one may employ such protocols to investigate the nature of the control-driven CSE.

The episodic retrieval account of the CSE

An emerging view of the control-driven CSE is the episodic retrieval account (Dignath, Johannsen, Hommel, & Kiesel, 2019; Egner, 2014; Spapé & Hommel, 2008; Weissman et al., 2016). Here, participants form an episodic memory of each trial that includes information about the stimuli and responses that occurred, various relationships between these stimuli and responses (e.g., congruent or incongruent), salient context-defining features (e.g., the sensory modality in which the stimuli appeared), and the task sets (e.g., stimulus-response mappings) that were employed. At the start of the next trial, repeating any of these features, and/or the local temporal context acting on its own, cues the retrieval of a memory of the previous trial. This biases control processes to employ the same control settings in the current trial, which leads to a smaller congruency effect after incongruent trials relative to congruent trials (i.e., a CSE).

A number of findings support the episodic retrieval account. First, stimulus and response repetitions (vs. alternations) in consecutive trials increase the CSE (Hommel et al., 2004; Mayr et al., 2003; Weissman et al., 2016). Second, repetitions of trial congruency (vs. alternations) engender a CSE even without stimulus and response repetitions (Jiménez & Méndez, 2014; Kim & Cho, 2014; Schmidt & Weissman, 2014). Third, the CSE is larger when context-defining features repeat (vs. alternate) in consecutive trials including the perceptual format in which the task stimuli appear (Dignath et al., 2019), the sensory modality in which the task stimuli appear (Hazeltine, Lightman, Schwarb, & Schumacher, 2011; Kreutzfeldt, Stephan, Willmes, & Koch, 2016; Yang et al., 2017), and the irrelevant voice gender with which an auditory distractor is spoken (Spapé & Hommel, 2008). Fourth, the CSE is larger when participants repeat the same task in consecutive trials than when they switch tasks (Kiesel, Kunde, & Hoffmann, 2006).

However, there is disagreement regarding certain aspects of the episodic retrieval view. More specifically, there are competing, yet complementary explanations of why changing a context-defining feature across consecutive trials reduces the CSE. We describe these explanations in greater detail in the next section.

Before proceeding, however, it is important to explain how we use four terms. First, we use the term “task” to refer to a representation of the cognitive processes and stimulus-response (S-R) mappings that enable participants to respond appropriately to relevant stimuli (Rogers & Monsell, 1995). Second, we posit that a “task set” is created when participants form an intention to perform a task, which involves actively maintaining that task in working memory (Rogers & Monsell, 1995). Third, as we described earlier, we use the term "context" to refer to salient perceptual aspects of the stimuli, such as their perceptual format (Dignath et al., 2019), sensory modality (Hazeltine et al., 2011), or the irrelevant voice gender in which an auditory distractor is spoken (Spapé & Hommel, 2008). Fourth, we define an “episode” as the temporal frame within which a participant perceives a trial to occur (Egner, 2014; Spapé & Hommel, 2008).

The attentional reset and task set hypotheses

The *attentional reset hypothesis* posits that changing a context-defining feature reduces or eliminates the CSE by resetting attentional processes (Kreutzfeldt et al., 2016). In this view, a sudden change to a context-defining feature indicates that the episodic context has changed (i.e., that a new episode has begun) and, consequently, that the control settings from the previous trial are no longer appropriate. This resets attentional processes, which eliminates the CSE. Critically, changing a context-defining feature triggers this reset of attentional processes even when the rest of the task stays the same. Thus, it is simply the change to this feature that reduces the CSE.

The *task set hypothesis* more narrowly defines the circumstances under which changing a context-defining feature reduces the CSE than the attentional reset hypothesis. Specifically, this hypothesis posits that changing a context-defining feature reduces the CSE only when the updated contextual feature is linked to a different task set (Hazeltine et al., 2011; Schumacher & Hazeltine, 2016). According to this view, control processes use different context-defining features (e.g., different sensory modalities) to assign different subsets of the task stimuli (e.g., visual vs. auditory stimuli) to different tasks (e.g., a visual task and an auditory task). When participants subsequently intend to perform one of these tasks, they must form a unique task set. Participants can employ the task set from the previous trial if the context-defining feature (e.g., sensory modality) from that trial repeats. In contrast, they must employ a different task set if the context-defining feature from the previous trial switches. Critically, switching to a different task set disrupts the ability to retrieve an episodic memory of the previous trial, because participants perceive that trial as part of a different event or episode (Hazeltine et al., 2011; Spapé & Hommel, 2008). Thus, switching between different task sets reduces the CSE.

Unlike the attentional reset hypothesis, the task set hypothesis does not always predict that changing a context-defining feature will reduce the CSE. This is because control processes may not always be able to employ context-defining features to divide a complex task into two simpler tasks. For example, some task structures may not allow control processes to assign stimuli in different sensory modalities to different tasks (e.g., Hazeltine et al., 2011, Experiment 4). In these situations, switching modalities should not reduce the CSE, because both modalities are associated with the same task and, hence, the same task set. In short, unlike the attentional reset hypothesis, the task set hypothesis posits that changing a context-defining feature should reduce the CSE only when the updated context-defining feature is associated with a different task set. Thus, relative to the attentional reset hypothesis, the task set hypothesis is more specific with regard to *when* changing a context-defining feature reduces the CSE.

The present study

In the present study, we employed a confound-minimized, prime-probe task to distinguish between the attentional reset and task set hypotheses. In Experiment 1, we sought to verify that the sensory modality in which task stimuli appear serves as a boundary for the CSE in the absence of feature integration and contingency learning confounds. Since, to our knowledge, all prior studies investigating modality-specific CSEs contained at least one of these confounds, we reasoned that observing such a result would provide a crucial starting point for distinguishing between the attentional reset and task set hypotheses. We observed this result in Experiment 1. Thus, we sought to distinguish between the attentional reset and task set hypotheses in Experiment 2. Finally, after obtaining support for the task set hypothesis, we investigated how modality-specific task sets form in Experiment 3.

Experiment 1

The goal of Experiment 1 was to verify that changing a context-defining feature can reduce the CSE even in the absence of feature integration and contingency learning confounds (Schmidt & Weissman, 2014; Weissman et al., 2015; Weissman et al., 2014). To this end, we conceptually replicated the task design from Experiment 1 of Hazeltine et al. (2011), wherein the common sensory modality in which both the distractor and target appear (visual or auditory) either repeats or switches in consecutive trials[6]. Hazeltine et al. (2011) observed a CSE when the sensory modality repeated across consecutive trials but not when it switched, consistent with a modality-specific boundary for the CSE. We reasoned that observing this outcome without the typical confounds would provide a starting point for distinguishing between the attentional reset and task set hypotheses in Experiment 2.

Both the task set hypothesis and the attentional reset hypothesis predict this outcome. The task set hypothesis predicts it because the distractor's modality always cues the target's modality. Participants can therefore use the distractor's modality to categorize every trial as "visual" or "auditory". This should allow control processes to assign visual and auditory stimuli to distinct visual and auditory tasks and, consequently, lead to the formation of modality-specific task sets (Schumacher & Hazeltine, 2016). Therefore, changing the modality in which task stimuli appear should disrupt the ability to retrieve an episodic memory of the previous trial and thereby reduce or eliminate the CSE (Hazeltine et al., 2011; Spapé & Hommel, 2008). In contrast, the attentional reset hypothesis predicts this outcome because it posits that any change in episodic context

---

[6] As noted by Spence (2010), stimuli in different modalities often appear at different spatial locations (e.g., on a computer screen or in headphones). This confound does not impact our present predictions, however, because both spatial location and sensory modality are context-defining features. Thus, the attentional reset and task set hypotheses both predict that changing one or both of these contextual features will reduce the CSE in Experiment 1.

signals that previous trial control settings are no longer appropriate. This leads to a reset of attentional processes, which eliminates the CSE (Kreutzfeldt et al., 2016).

Methods

Participants

To determine the sample size needed to observe a modality-specific boundary for the CSE, we conducted a power analysis in G*Power 3.1.9.2 (Faul, Erdfelder, Lang, & Buchner, 2007). To do so, we combined an alpha (α) of 0.05 with the effect size ($\boxed{\eta_p^2 =}$0.372) associated with the three-way interaction that Hazeltine et al. (2011, Experiment 1) observed among modality condition (repeat, switch), previous trial congruency (congruent, incongruent), and current trial congruency (congruent, incongruent). The results indicated that collecting usable data from 32 participants would provide over 99% power to observe this interaction.

Thirty-seven undergraduates from the University of Michigan participated in Experiment 1 for course credit, five of whom were excluded based on our a priori exclusion criteria. More specifically, we excluded one who self-reported neurological disorders and/or the use of psychoactive medications, one who experienced a computer malfunction during the task, and three who performed the task with less than 70% accuracy. The remaining 32 participants (27 male, 5 female; 27 right-handed, 5 left-handed; age range, 18-20 years; mean age, 18.56 years; standard deviation of age, 0.72 years) reported normal or corrected-to-normal vision with no history of neuropsychiatric illness, seizures, or head trauma. The University of Michigan's Behavioral Sciences Internal Review Board approved all experimental procedures.

Stimuli and Apparatus

We used the Psychophysics Toolbox running on a Windows PC to present the stimuli and to record participants' responses (Brainard, 1997). Three types of stimuli were presented: (1) a 2 s fixation cross (0.8° x 0.8°), (2) spoken auditory distractor and target words that indicated left, right, up, or down (i.e., "Left": 63 dB, "Right": 64 dB, "Up": 64 dB, or "Down": 68 dB), and (3) visual distractor and target words that indicated left, right, up, or down (i.e., "Left": 5.6° x 1.7°, "Right": 7.0° x 1.7°, "Up": 3.0° x 1.7°, or "Down": 7.3° x 1.7°). To create the auditory stimuli, we used a MATLAB script that converts written text to speech (https://www.mathworks.com/matlabcentral/fileexchange/18091-text-to-speech). All visual stimuli were presented on a black background at a viewing distance of approximately 55 cm. We collected responses using a standard QWERTY keyboard.

Experimental Design

In each 2.5-second trial, there were four sequential events: a distractor (duration, 250 ms), a blank screen (duration, 250 ms), a target (duration, 250 ms), and a second blank screen (duration, 1750 ms). We created 16 distractor-target pairs using the visual and auditory stimuli described earlier. There were eight visual distractor-target pairs and eight auditory distractor-target pairs. Each set of eight included four congruent pairs (left-left, right-right, up-up, and

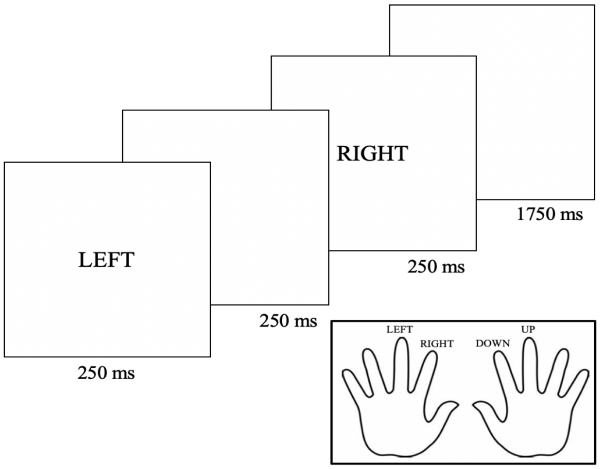


Figure 4-1. The prime-probe task used in Experiment 1
In each trial, a distractor word preceded a target word (the figure illustrates one trial). Trials contained either visual stimuli (shown above) or auditory stimuli. Visual stimuli appeared in white on a black background. Auditory stimuli were delivered via headphones. The number beneath each box indicates the length of the corresponding trial component in milliseconds (ms). The hands indicate the stimulus-response mapping.

down-down) and four incongruent pairs (left-right, right-left, up-down, and down-up). To avoid feature integration confounds (Hommel et al., 2004; Mayr et al., 2003), we presented pairs composed of left-right stimuli in the odd trials of each block and pairs composed of up-down stimuli in the even trials of each block. To avoid contingency learning confounds (Schmidt & De Houwer, 2011), we presented all distractor-target pairs equally often in every block. To ensure that we could compare the CSE when the modality changed to the CSE when the modality repeated, we employed a first-order counterbalanced trial sequence in each block. In this sequence, four trial types (visual congruent, visual incongruent, auditory congruent, and auditory incongruent) preceded and followed each other equally often, separately for odd and even trials. One of the sixteen resulting sequential trial types appeared one fewer time than the others, because the first trial in the block was not preceded by a trial. However, the underrepresented sequential trial type varied randomly across blocks, because we employed a unique trial sequence in every block.

Procedure

After they arrived at the laboratory, participants provided informed written consent, filled out a brief screening form, and performed the prime-probe task. We instructed participants to respond as quickly and accurately as possible in each trial by pressing the F, G, J, or N key to indicate whether the target word was left, right, up, or down, respectively (Fig. 1). Each time the participant responded incorrectly or failed to respond within 1.25 s of target onset, the word "Error" appeared centrally for 200 ms. A 64-trial practice block was followed by ten 64-trial test blocks. Each block started and ended with a 2 s fixation cross.

Data Analyses

In our analyses of mean response time (RT), we excluded (1) practice trials, (2) the first trial of each block, (3) outliers in which reaction time was at least three standard deviations from the condition mean, (4) trials with omitted or incorrect responses, and (5) trials immediately following incorrect or omission trials. These trial types were also excluded in our analyses mean error rate (ER), with the exception of the fourth trial type because error rate was the dependent measure. On average, 6.2% of the trials were errors and 1.0% were outliers.

Following these exclusions, we calculated mean RT and mean ER for each trial type. We then conducted separate repeated-measures ANOVAs on mean RT and mean ER. In each ANOVA, the three within-participants factors were modality condition (repeat, switch), previous trial congruency (congruent, incongruent), and current trial congruency (congruent, incongruent). We also calculated the simple effects that were associated with the three-way interaction.

In addition to these ANOVAs, we computed repeated-measures Bayesian ANOVAs using JASP (JASP Team, 2019). Specifically, we computed Bayes factors in favor of the alternative hypothesis (i.e., BF_{10}) and in favor of the null hypothesis (i.e., BF_{01}), separately for each main effect and interaction. To do so, we compared a separate model of (a) each main effect to a null model with no main effects or interactions, (b) each two-way interaction to a null model with all main effects and no interactions and (c) the three-way interaction to a null model with all main effects and all two-way interactions. To interpret these Bayes factors, we used Raftery's (1995) classification theme, wherein a Bayes factor ranging from (a) 1-3 is considered weak evidence, (b) 3-20 is considered positive evidence, (c) 20-150 is considered strong evidence, and (d) any value greater than 150 is considered extremely strong evidence. The results of each ANOVA are listed in Table 1.

Results

Mean RT

We observed three significant main effects. First, we observed a main effect of modality condition, *F*(1,31) = 12.63, *p* = 0.001, η_p^2 = 0.29, BF$_{10}$ = 0.44, BF$_{01}$ = 2.29, because mean RT was longer in modality switch trials (524 ms) than in modality repeat trials (516 ms). Second, we observed a main effect of current trial congruency (i.e., a congruency effect), *F*(1,31) = 55.36, *p* < 0.001, η_p^2 = 0.64, BF$_{10}$ = 3.63 x 10^{32}, BF$_{01}$ = 2.77 x 10^{-33}, because mean RT was longer in incongruent trials (544 ms) than in congruent trials (496 ms). Third, we observed a main effect of previous trial congruency, *F*(1,31) = 7.22, *p* = 0.011, η_p^2 = 0.19, BF$_{10}$ = 0.30, BF$_{01}$ = 3.33: mean RT was longer after incongruent trials (523 ms) than after congruent trials (517 ms), consistent with "post-conflict" slowing (Ullsperger et al., 2005).

We also observed a significant two-way interaction. Specifically, we observed an interaction between previous trial congruency and current trial congruency (i.e., a CSE), *F*(1,31) = 18.55, *p* < 0.001, η_p^2 = 0.37, BF$_{10}$ = 9.51, BF$_{01}$ = 0.10. As expected, the congruency effect was smaller after incongruent trials (39 ms) than after congruent trials (58 ms).

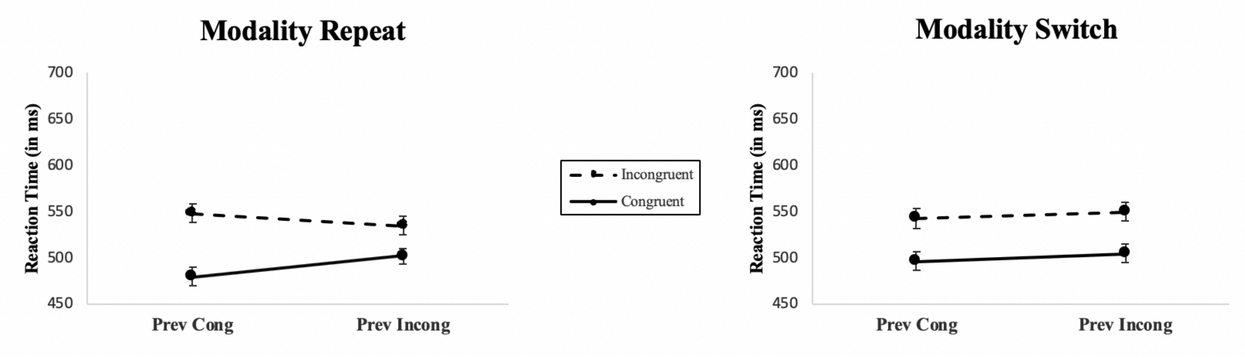


Figure 4-2. The CSE in each of the two main trial types of Experiment 1

Previous trial congruency is indicated on the x-axis (Prev Cong: previous congruent trial; Prev Incong: previous incongruent trial). Current trial congruency is indicated by line type (Incong: dashed line; Cong: black line). Reaction Time (in ms) is indicated on the y-axis. Error bars indicate ±1 standard error of the mean.

Finally, we observed a significant three-way interaction. In line with both the task set and attentional reset hypotheses, we observed an interaction among modality condition, previous trial congruency, and current trial congruency, $F(1,31) = 23.20$, $p < 0.001$, $\boxed{\eta_p^2} = 0.43$, $BF_{10} = 7.49$, $BF_{01} = 0.16$[7]. As expected, the CSE was larger in modality repeat trials (35 ms; $F(1,31) = 36.01$, $p < 0.001$, $\boxed{\eta_p^2} = 0.54$, $BF_{10} = 74.54$, $BF_{01} = 0.013$) than in modality switch trials (2 ms; $F(1,31) < 1$, $BF_{10} = 0.27$, $BF_{01} = 3.75$) (Fig. 2). No other effects were significant (see Table 1).

Mean ER

We observed three significant main effects. First, we observed a main effect of modality condition, $F(1,31) = 9.49$, $p = 0.004$, $\boxed{\eta_p^2} = 0.23$, $BF_{10} = 12.41$, $BF_{01} = 0.08$, because mean ER was lower when the modality switched (2.5%) than when it repeated (3.7%). Second, we observed a main effect of previous trial congruency, $F(1,31) = 8.03$, $p = 0.008$, $\boxed{\eta_p^2} = 0.21$, $BF_{10} = 4.83$, $BF_{01} = 0.21$, because mean ER was lower after incongruent trials (2.5%) than after congruent trials (3.6%). Along with the mean RT data, this effect suggests that participants traded speed for accuracy after incongruent relative to congruent trials, consistent with post-conflict slowing (e.g., Ullsperger et al., 2005). Third, as in the mean RT data, we observed a main effect of current trial congruency, $F(1,31) = 12.03$, $p = 0.002$, $\boxed{\eta_p^2} = 0.28$, $BF_{10} = 2.42 \times 10^3$,

---

[7] Interestingly, exploratory analyses indicated a significant interaction among current trial modality, previous trial congruency, and current trial congruency, $F(1, 31) = 5.79$, $p = 0.02$, $\boxed{\eta_p^2} = 0.16$, $BF_{10} = 13.95$, $BF_{01} = 0.07$, because the CSE was larger in visual modality trials (26 ms) than in auditory modality trials (10 ms). However, current trial modality did not influence the interaction among modality condition, previous trial congruency, and current trial congruency ($F < 1$, $BF_{10} = 0.26$, $BF_{01} = 3.84$). In other words, the modality of the current trial did not influence the degree to which the CSE was modality-specific.

$BF_{01} = 4.13 \times 10^{-4}$, as mean ER was higher in incongruent trials (3.9%) than in congruent trials (2.2%).

We also observed three significant interactions. First, we observed a two-way interaction between modality condition and current trial congruency, $F(1,31) = 9.03$, $p = 0.005$, $\boxed{\eta_p^2} = 0.23$, $BF_{10} = 2.85$, $BF_{01} = 0.35$: the congruency effect was larger in modality repeat trials (2.6%) than in modality switch trials (0.9%). Second, we observed a two-way interaction between previous trial congruency and current trial congruency, $F(1,31) = 5.56$, $p = 0.025$, $\boxed{\eta_p^2} = 0.15$, $BF_{10} = 0.94$, $BF_{01} = 1.06$, because the congruency effect was smaller after incongruent trials (1.1%) than after congruent trials (2.4%). Third, as in the mean RT data, we observed a three-way interaction among modality condition, previous trial congruency, and current trial congruency, $F(1,31) = 5.18$, $p = 0.030$, $\boxed{\eta_p^2} = 0.14$, $BF_{10} = 0.95$, $BF_{01} = 1.06$, because the CSE was larger in modality repeat trials (2.6%; $F(1,31) = 9.93$, $p = 0.004$, $\boxed{\eta_p^2} = 0.24$, $BF_{10} = 3.70$, $BF_{01} = 0.27$) than in modality switch trials (0.1%; $F(1,31) < 1$, $BF_{10} = 0.25$, $BF_{01} = 4.01$). No other effects were significant (see Table 1).

Table 4-1. Statistics for repeated-measures and Bayesian ANOVAs in Exp. 1

Dependent	Effect	Repeated-measures			Bayesian	
		F	*P*	η_p^2	BF_{10}	BF_{01}
Mean RT	Modality Condition	12.63	.001	.29	0.44	2.29
	Previous Trial Congruency	7.22	.01	.19	0.30	3.33
	Current Trial Congruency	55.36	<.001	.64	3.63×10^{32}	2.77×10^{-33}
	Modality Condition x Previous Trial Congruency	0.71	.41	.02	0.21	4.94
	Modality Condition x Current Trial Congruency	1.80	.19	.06	0.24	4.22
	Previous Trial Congruency x Current Trial Congruency	18.55	<.001	.37	9.51	0.10
	3-way interaction	23.20	<.001	.43	7.49	0.16
Mean ER	Modality Condition	9.45	.004	.23	12.41	0.08
	Previous Trial Congruency	8.03	.008	.21	4.83	0.21
	Current Trial Congruency	12.03	.002	.28	2.42×10^{3}	4.13×10^{-4}
	Modality Condition x Previous Trial Congruency	1.05	.31	.03	0.25	4.00
	Modality Condition x Current Trial Congruency	9.03	.005	.23	2.85	0.35
	Previous Trial Congruency x Current Trial Congruency	5.56	.025	.15	0.94	1.06
	3-way interaction	5.18	.030	.14	0.95	1.06

Discussion

Consistent with control-based accounts of the CSE, we found that changing the sensory modality in which task stimuli appear eliminates the CSE in the absence of feature integration and contingency learning confounds. This outcome bolsters prior data suggesting modality-specific boundaries for cognitive control processes (Hazeltine et al., 2011; Kreutzfeldt et al., 2016; Yang et al., 2017). It also serves as a crucial starting point for investigating whether modality-specific CSEs are more consistent with the attentional reset or task set hypothesis.

Experiment 2

The goal of Experiment 2 was to distinguish between the attentional reset and task set hypotheses. To this end, we independently varied the sensory modality (i.e., visual or auditory) in which the distractor and target appeared across trials. Thus, in addition to the "all visual" and "all auditory" trials of Experiment 1, we included "mixed modality" trials wherein an auditory distractor was followed by a visual target or a visual distractor was followed by an auditory target. As a result of including pure and mixed modality trials with equal frequency, the modality of the distractor in a given trial was not predictive of the modality of the target in that trial. Critically, the inability to classify each trial as either "visual" or "auditory" should prevent control processes from assigning stimuli in different sensory modalities to different tasks and, consequently, preclude the formation of modality-specific task sets. Therefore, in this design, the task set hypothesis predicts no reduction of the CSE when both the distractor's modality and the target's modality switch in consecutive trials, relative to when they repeat, because both modalities are associated with the same task set. In contrast, the attentional reset hypothesis continues to predict exactly such a reduction. In this view, any salient change in episodic context (e.g., a change to the sensory modality in which task stimuli appear) signals that previous-trial control settings are no longer appropriate (e.g., Kreutzfeldt et al., 2016), regardless of how this contextual change relates to the overarching task structure. Thus, the CSE should always be reduced when both the distractor's modality and the target's modality switch in consecutive trials relative to when they repeat.

Methods

Participants

As in Experiment 1, we conducted a power analysis in G*Power 3.1.9.2 (Faul et al., 2007) to estimate the sample size needed to observe the critical three-way interaction among modality condition, previous trial congruency, and current trial congruency. To do so, we used the effect size ($\boxed{\eta_p^2 =}$0.428) for this interaction from Experiment 1 of the present study (α = 0.05). The results indicated that 32 participants would provide over 99% power to observe this interaction in the present experiment.

Thirty-six undergraduates from the University of Michigan participated for course credit, four of whom were excluded based on the criteria in Experiment 1: one who did not complete the task and three who performed the task with less than 70% accuracy. The remaining 32 participants (all male; 27 right-handed, 5 left-handed; age range, 17-21 years; mean age, 18.59 years; standard deviation of age, 0.84 years) reported normal or corrected-to-normal vision with no history of neuropsychiatric illness, seizures, or head trauma. The University of Michigan's Behavioral Sciences Internal Review Board approved all experimental procedures.

Stimuli and Apparatus

The stimuli and apparatus were identical to those in Experiment 1.

Experimental Design

The experimental design was similar to that in Experiment 1 with two exceptions. First, we created 32 (rather than 16) distractor-target pairs involving the visual and auditory word stimuli. This allowed us to independently vary the sensory modality of the distractor (visual, auditory) and the target (visual, auditory). In total, there were eight visual-visual, eight auditory-auditory, eight visual-auditory, and eight auditory-visual distractor-target pairs. Second, we employed a first-order counterbalanced trial sequence in each block that involved four "same modality" trial types (visual congruent, visual incongruent, auditory congruent, and auditory incongruent) and four "mixed modality" trial types (visual-auditory congruent, visual-auditory incongruent, auditory-visual congruent, and auditory-visual incongruent). That is, there were eight trial types in total rather than four trial types as in Experiment 1. These trial types preceded and followed each other equally often, separately for odd and even trials.

Procedure

The procedure was the same as that in Experiment 1 with one exception. Specifically, the experiment consisted of a 64-trial practice block followed by five 128-trial (versus ten 64-trial) test blocks. We inserted a rest break halfway through each block to equate the number and frequency of rest breaks to those in Experiment 1.

Data Analyses

The data analysis was identical to that in Experiment 1 with two exceptions. First, we considered modality repeat trials as consecutive trials wherein both the distractor and the target modalities independently repeated (e.g., a visual-visual trial followed by a visual-visual trial or an auditory-visual trial followed by an auditory-visual trial). Second, we considered modality switch trials as consecutive trials wherein both the distractor and target modalities independently switched (e.g., a visual-visual trial followed by an auditory-auditory trial or an auditory-visual trial followed by a visual-auditory trial). These changes allowed us to include more trials in the analysis. Therefore, we reasoned that they would likely increase statistical power. On average, 7.5% of the trials were errors and 0.6% were outliers.

Results

Mean RT

We observed two significant main effects in the mean RT data. First, we observed a main effect of modality condition, $F(1,31) = 5.17$, $p = 0.030$, $\boxed{\eta_p^2} = 0.14$, $BF_{10} = 0.31$, $BF_{01} = 3.25$, because mean RT was longer in modality switch trials (598 ms) than in modality repeat trials (589 ms)[8]. Second, we observed a main effect of current trial congruency, $F(1,31) = 38.68$, $p < 0.001$, $\boxed{\eta_p^2} = 0.56$, $BF_{10} = 1.22 \times 10^{22}$, $BF_{01} = 8.20 \times 10^{-23}$, because mean RT was longer in incongruent trials (622 ms) than in congruent trials (565 ms).

We also observed a significant two-way interaction. Specifically, we observed an interaction between previous trial congruency and current trial congruency, $F(1,31) = 37.96$, $p <$

---

[8] Notably, the effect of modality condition – an index of modality task-switch costs – was rather small in both Experiments 1 and 2 (~8 ms). One possible explanation is that we employed a relatively long (500 ms) interval between the distractor and target. Indeed, long cuing intervals are associated with reduced modality task-switch costs (e.g., Lukas, Philipp & Koch, 2010). Thus, the degree of cross-modal competition may have been smaller in the present study than in studies wherein researchers employed shorter (e.g., 200 ms) cueing intervals (e.g., Kreutzfeldt et al., 2016).

0.001, $\boxed{\eta_p^2}$ = 0.55, BF_{10} = 48.92, BF_{01} = 0.02. As expected, the congruency effect was smaller after incongruent trials (41 ms) than after congruent trials (75 ms).

Crucially, unlike in Experiment 1, we did not observe a significant three-way interaction among modality condition, previous trial congruency, and current trial congruency, $F(1,31) < 1$, BF_{10} = 0.26, BF_{01} = 3.90 (Fig. 3). Rather, the CSE was robust in both modality repeat trials (37 ms; $F(1,31) = 20.57$, $p < 0.001$, $\boxed{\eta_p^2}$ = 0.40, BF_{10} = 6.26, BF_{01} = 0.16) and modality switch trials (32 ms; $F(1, 31) = 14.88$, $p < 0.001$, $\boxed{\eta_p^2}$ = 0.32, BF_{10} = 1.48, BF_{01} = 0.67). Moreover, exploratory analyses revealed that (a) this interaction was not significant even when we restricted the analysis to “all visual” and “all auditory” trials ($p > 0.20$, BF_{10} = 0.35, BF_{01} = 2.83) and (b) independent changes to the distractor’s modality, the target’s modality, or both did not influence the size of the CSE (all p-values ≥ 0.10; all BF_{10} values < 0.30; all BF_{01} values > 3.0). No other effects were significant (see Table 2).

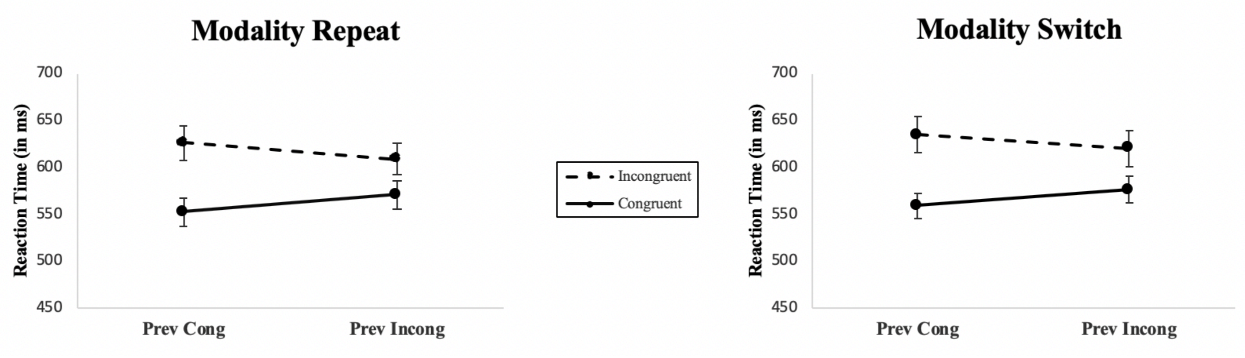


Figure 4-3. The CSE in each of the two main trial types of Experiment 2

Previous trial congruency is indicated on the x-axis (Prev Cong: previous congruent trial; Prev Incong: previous incongruent trial). Current trial congruency is indicated by line type (Incong: dashed line; Cong: black line). Reaction Time (in ms) is indicated on the y-axis. Error bars indicate ±1 standard error of the mean.

<u>Mean ER</u>

We observed two significant main effects in the mean ER data. First, we observed a main effect of previous trial congruency, $F(1,31) = 19.08$, $p < 0.001$, $\eta_p^2 = 0.38$, $\mathrm{BF}_{10} = 15.54$, $\mathrm{BF}_{01} = 0.064$, because mean ER was lower after incongruent trials (3.5%) than after congruent trials (5.4%). Second, as in the mean RT data, we observed a main effect of current trial congruency, $F(1,31) = 23.76$, $p < 0.001$, $\eta_p^2 = 0.43$, $\mathrm{BF}_{10} = 2.70 \times 10^{11}$, $\mathrm{BF}_{01} = 3.71 \times 10^{-12}$: mean ER was higher in incongruent trials (6.6%) than in congruent trials (2.3%).

We also observed a significant two-way interaction. Specifically, we observed an interaction between previous trial congruency and current trial congruency, $F(1,31) = 21.79$, $p < 0.001$, $\eta_p^2 = 0.41$, $\mathrm{BF}_{10} = 115.49$, $\mathrm{BF}_{01} = 0.01$, because the congruency effect was smaller after incongruent trials (2.4%) than after congruent trials (6.2%). Finally, although it did not reach conventional levels of significance, we observed a trend toward an interaction between modality condition and current trial congruency, $F(1,31) = 4.04$, $p = 0.053$, $\eta_p^2 = 0.12$, $\mathrm{BF}_{10} = 0.35$, $\mathrm{BF}_{01} = 2.86$, because the congruency effect was smaller in modality repeat trials (3.7%) than in modality switch trials (5.0%).

As in the mean RT analyses, we did not observe a significant three-way interaction among modality condition, previous trial congruency, and current trial congruency, $p > 0.30$, $\mathrm{BF}_{10} = 0.32$, $\mathrm{BF}_{01} = 3.18$. However, exploratory analyses revealed a pair of interesting findings. First, the three-way interaction was not significant even when we restricted the analysis to "all visual" and "all auditory" trials ($F < 1$, $\mathrm{BF}_{10} = 0.27$, $\mathrm{BF}_{01} = 3.67$). Second, although switching only the distractor's modality or only the target's modality did not influence the size of the CSE (all F-values < 1; all BF_{10} values < 0.55; all BF_{01} values > 1.5), switching both of these modalities did exert an influence, $F(1,31) = 12.63$, $p = 0.001$, $\eta_p^2 = 0.29$, $\mathrm{BF}_{10} = 3.11$, $\mathrm{BF}_{01} = 0.32$. Nonetheless, as in the mean RT data, the CSE was significant regardless of whether both modalities repeated (4.7%; $F(1,31) = 13.41$, $p < 0.001$, $\eta_p^2 = 0.30$, $BF_{10} = 3.30$, $BF_{01} = 0.30$) or switched (2.8%; $F(1,31) = 6.71$, $p = 0.015$, $\eta_p^2 = 0.18$, $BF_{10} = 1.11$, $BF_{01} = 0.90$) across consecutive trials. No other effects were significant (see Table 2).

Table 4-2. Statistics for repeated-measures and Bayesian ANOVAs in Exp. 2

Dependent	Effect	Repeated-measures			Bayesian	
		F	p	η_p^2	BF_{10}	BF_{01}
Mean RT	Modality Condition	5.17	.030	.14	0.31	3.25
	Previous Trial Congruency	.06	.82	.00	0.13	7.47
	Current Trial Congruency	38.68	<.001	.56	1.22×10^{22}	8.20×10^{-23}
	Modality Condition x Previous Trial Congruency	.00	.98	.00	0.18	5.46
	Modality Condition x Current Trial Congruency	.33	.57	.01	0.54	1.86
	Previous Trial Congruency x Current Trial Congruency	37.96	<.001	.55	48.92	0.02
	3-way interaction	.19	.66	.01	0.26	3.90
Mean ER	Modality Condition	.31	.58	.01	0.15	6.61
	Previous Trial Congruency	19.08	< .001	.38	15.54	0.06
	Current Trial Congruency	23.76	< .001	.43	2.70×10^{11}	3.71×10^{-12}
	Modality Condition x Previous Trial Congruency	.50	.49	.02	0.19	5.38
	Modality Condition x Current Trial Congruency	4.04	.053	.12	0.35	2.86
	Previous Trial Congruency x Current Trial Congruency	21.79	< .001	.41	115.49	0.01
	3-way interaction	1.01	.32	.03	0.32	3.18

Exploratory Across-Experiment Analysis

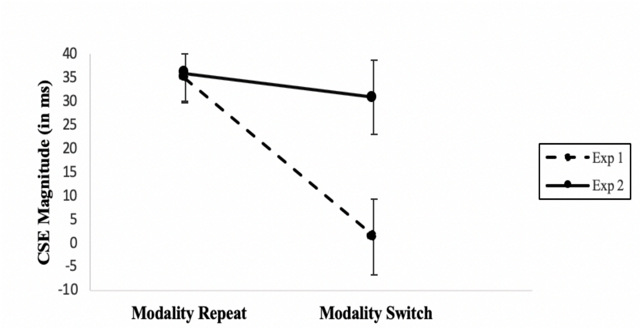


Figure 4-4. The CSE in each of the two main trial types of Experiments 1 and 2

Condition (repeat, switch) is indicated on the x-axis. Experiment is indicated by line type (Exp. 1: dashed line; Exp. 2: black line). CSE Magnitude (in ms) is indicated on the y-axis. Error bars indicate ±1 standard error of the mean.

Consistent with the task set hypothesis, our findings suggest that changing the sensory modality in which the task stimuli appear reduces the CSE more when participants form modality-specific task sets (Experiment 1) than when they do not (Experiment 2). As a further test of this hypothesis, we explored whether the reduction of the CSE associated with switching (versus repeating) sensory modalities was greater in Experiment 1 than in Experiment 2. To do so, we conducted separate mixed ANOVAs on mean RT and mean ER wherein Experiment (1, 2) served as the across-participants factor and modality condition (repeat, switch), previous trial congruency (congruent, incongruent), and current trial congruency (repeat, switch) served as the within-participants factors. The mean RT analysis revealed (1) a significant interaction among Experiment, previous trial congruency, and current trial congruency, $F(1,62) = 4.13$, $p = 0.047$, $\boxed{\eta_p^2} = 0.06$, $BF_{10} = 0.46$, $BF_{01} = 2.17$ and (2) a trending interaction among these four factors, $F(1,62) = 3.73$, $p = 0.058$, $\boxed{\eta_p^2} = 0.06$, $BF_{10} = 0.45$, $BF_{01} = 2.24$. The first finding likely reflects a smaller CSE in Experiment 1 (19 ms) than in Experiment 2 (34 ms). The second finding likely reflects a larger numerical difference in CSE magnitude between modality repeat trials and modality switch trials in Experiment 1 (32 ms) than in Experiment 2 (5 ms). Finally, similar to what we observed in the mean RT analysis, the mean ER analysis revealed an interaction among Experiment, previous trial congruency, and current trial congruency, $F(1,62) = 5.19$, $p = 0.026$, $\boxed{\eta_p^2} = 0.08$, $BF_{10} = 0.46$, $BF_{01} = 2.17$. However, the four-way interaction was not significant $F(1,62) < 1$, $BF_{10} = 0.17$, $BF_{01} = 5.76$.

Discussion

In Experiment 2, we sought to distinguish between the attentional reset and task set hypotheses by independently varying the sensory modalities (i.e., visual and auditory) of the distractor and the target. As described earlier, the task set hypothesis posits that the inability to classify each trial as either "visual" or "auditory" should prevent participants from using the sensory modality in which task stimuli appear to form modality-specific task sets (Schumacher & Hazeltine, 2016). Thus, the task set hypothesis predicts a CSE regardless of whether the sensory modality of (a) the distractor and (b) the target both repeat or both switch in consecutive trials. In contrast, the attentional reset hypothesis predicts a CSE only when the sensory modality of each stimulus repeats, because any change to the episodic context signals that previous trial control settings are no longer relevant (Kreutzfeldt et al., 2016). Consistent with the task set hypothesis, but not with the attentional reset hypothesis, the interaction among modality condition, previous trial congruency, and current trial congruency was (a) not significant and (b) more consistent with the null hypothesis ($BF_{01} = 3.90$) than with the alternative hypothesis ($BF_{10} = 0.26$). Furthermore, the CSE was highly significant in both modality repeat trials and modality switch trials (see the mean RT data). Together, these findings provide novel support for the task set hypothesis.

Experiment 3

While Experiments 1 and 2 provide strong support for the task set hypothesis, they do not provide insight into the mechanism(s) that form a modality-specific task set. Recent findings from Kreutzfeldt et al. (2016), however, suggest that orienting attention to the sensory modality in which the distractor is presented before the target appears may be one such mechanism. In this study, participants were cued on a trial-by-trial basis to identify either the visual or the auditory aspect of an upcoming audio-visual stimulus. To respond both quickly and correctly, participants needed to orient their attention to the cued sensory modality prior to display onset. Most important for present purposes, changing the cued sensory modality from one trial to the next eliminated the CSE. This outcome suggests that orienting attention to a sensory modality (i.e., a context-defining feature) may lead to the formation of a modality-specific task set that serves as a subsequent boundary for the CSE.

Notably, this orienting hypothesis may explain the contrasting effects that we observed in Experiments 1 and 2. Indeed, in Experiment 1 the distractor's modality always predicted the modality of the upcoming target. Thus, orienting attention to the distractor's modality would be useful to prepare for the upcoming target. In Experiment 2, however, the distractor's modality did not systematically predict the target's modality. Thus, orienting to the distractor's modality would not be useful to prepare for the upcoming target. Critically, we observed modality-specific CSEs in Experiment 1 but not in Experiment 2. Analogous to the findings of Kreutzfeldt et al. (2016), this result suggests that orienting attention to a sensory modality may result in the formation of a modality-specific task set that serves as a subsequent boundary for the CSE.

In Experiment 3, we investigated whether the formation of modality-specific task sets is associated with orienting attention to the sensory modality in which the distractor appears. To this end, we modified the task design from Experiment 2 wherein the distractor's modality did not systematically predict the target's modality. In the modified design, the distractor's modality predicted the target's modality in 89.9% of trials (same modality trials) but not in the remaining 11.1% of trials (mixed modality trials). We reasoned that employing these percentages of same and mixed modality trials would encourage participants to orient attention to the distractor's modality to prepare for the upcoming target while still allowing us to include a minority of trials wherein the distractor and target appear in different modalities. Thus, we also reasoned that this design would allow us to measure whether participants orient attention to the distractor's modality via the *modality validity effect* (Turatto, Benso, Galfano, & Umilta, 2002). In the present study, this effect refers to longer mean RT when a target appears in a different modality than a preceding distractor (invalid trials) relative to the same modality (valid trials).

The orienting hypothesis makes two predictions. First, it predicts modality-specific CSEs coupled with a modality validity effect. Second, it predicts a larger modality validity effect than in Experiment 2, wherein the distractor's modality did not systematically predict the target's modality. Indeed, in the absence of such a systematic relationship in Experiment 2, orienting attention to the distractor's modality would not be useful to prepare for the upcoming target.

Methods

Participants

To estimate the sample size needed to investigate our predictions, we conducted a series of power analyses ($\alpha = 0.05$) in G*Power (Faul et al., 2007). The results indicated that 32 participants would provide over 99% power to observe a three-way interaction among modality condition, previous trial congruency, and current trial congruency ($\boxed{\eta_p^2 =}$0.428 from Experiment

1 of the present study). They also indicated that this sample size would provide over 99% power to observe (1) a modality validity effect (η_p^2 = 0.557 from Experiment 1 of Turatto et al., 2002) and (2) an interaction between Experiment (2, 3) and the modality validity effect ($\eta_p^2 =$ 0.362 from Experiment 2 of Turatto et al., 2002).

Thirty-four undergraduates from the University of Michigan participated in the experiment for course credit, two of whom were excluded based on the same criteria in Experiments 1 and 2 (both reported neurological disorders). The remaining 32 participants (8 male, 24 female; 28 right-handed, 4 left-handed; age range, 18-21 years; mean age, 18.53 years; standard deviation of age, 0.80 years) reported normal or corrected-to-normal vision with no history of neuropsychiatric illness, head trauma, or seizures. The University of Michigan's Behavioral Sciences Internal Review Board approved all experimental procedures.

Stimuli and Apparatus

The stimuli and apparatus were identical to those in Experiment 2.

Experimental Design

The experimental design was identical to that in Experiment 2 with one exception. Specifically, we presented same modality trials 89.9% of the time (versus 50%) and mixed modality trials only 11.1% of the time (versus 50%). To accommodate this change, we created 18 (versus 32) distractor-target pairs. As in Experiment 2, there were 16 same-modality distractor-target pairs. However, there were only two (versus 16) mixed-modality distractor-target pairs. We created the trial sequence by counterbalancing the 16 same-modality and 2 mixed-modality trial types, separately for odd and even trials.

Procedure

The procedure was identical to that in Experiment 2 with one exception. The experiment consisted of a 54-trial (versus a 64-trial) practice block followed by a 648-trial test block (versus five 128-trial test blocks). We included a single 648-trial test block so that we could first-order counterbalance the 16 same-modality trial types and the 2 mixed-modality trial types separately for odd and even trials. To roughly equate the number and frequency of rest breaks to those in prior experiments, we inserted a rest break every 54 trials throughout the 648-trial test block.

Data Analyses

We conducted a series of analyses to investigate our predictions. To test our predictions for the modality validity effect, we conducted a mixed ANOVA. In this ANOVA, Experiment (2, 3) served as a between-participants factor, validity (valid, invalid) served as a within-participants factor, and mean RT served as the dependent measure. To test our predictions for the CSE, we conducted separate repeated-measures ANOVAs on mean RT and mean ER with three within-participants factors: modality condition (repeat, switch), previous trial congruency (congruent, incongruent), and current trial congruency (congruent, incongruent)[9]. On average, 4.1% of the trials were errors and 1.2% were outliers.

---

[9] We did not include validity (valid, invalid) as a fourth factor in the repeated-measures ANOVAs, because there were an insufficient number of trials in the conditions involving invalid trials.

Results

Mixed ANOVA to assess the modality validity effect

The analysis of the modality validity effect yielded two main results. First, as predicted, the modality validity effect was larger in Experiment 3 than in Experiment 2 ($F(1,62) = 19.86$, $p < 0.001$, $\eta_p^2 = 0.24$, $BF_{10} = 487.54$, $BF_{01} = 0.00$). Follow-up analyses of the simple effects of this interaction revealed a highly significant modality validity effect in Experiment 3 (44 ms; $F(1,62) = 33.15$, $p < 0.001$, $\eta_p^2 = 0.52$, $BF_{10} = 4.64 \times 10^{3}$, $BF_{01} = 2.15 \times 10^{-4}$) but not in Experiment 2 (0 ms; $F(1,62) < 1$, $BF_{10} = 0.25$, $BF_{01} = 3.99$). Both of these findings support the orienting hypothesis.

Repeated-measures ANOVAs to assess modality-specific CSEs

Mean RT

We observed two significant main effects. First, we observed a main effect of modality condition, $F(1,31) = 11.22$, $p = 0.002$, $\eta_p^2 = 0.27$, $BF_{10} = 3.07$, $BF_{01} = 0.33$, because mean RT was longer in modality switch trials (588 ms) than in modality repeat trials (576 ms). Second, we observed a main effect of current trial congruency, $F(1,31) = 36.77$, $p < 0.001$, $\eta_p^2 = 0.54$, $BF_{10} = 1.84 \times 10^{19}$, $BF_{01} = 5.44 \times 10^{-20}$, because mean RT was longer in incongruent trials (603 ms) than in congruent trials (560 ms). Finally, although it did not reach conventional levels of significance, there was a trend toward a main effect of previous trial congruency, $F(1,31) = 3.87$,

$p = 0.058$, $\eta_p^2 = 0.11$, $BF_{10} = 0.26$, $BF_{01} = 3.85$, because mean RT was longer after incongruent trials (584 ms) than after congruent trials (579 ms).

We also observed a significant two-way interaction. Specifically, we observed an interaction between previous trial congruency and current trial congruency, $F(1,31) = 4.84$, $p = 0.035$, $\eta_p^2 = 0.14$, $BF_{10} = 0.64$, $BF_{01} = 1.57$. As expected, the congruency effect was smaller after incongruent trials (37 ms) than after congruent trials (50 ms).

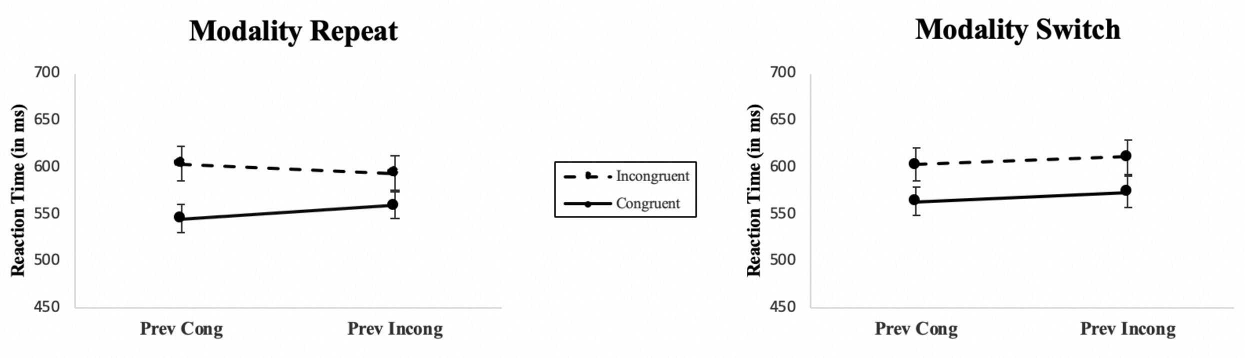


Figure 4-5. The CSE in each of the two main trial types of Experiment 3

Previous trial congruency is indicated on the x-axis (Prev Cong: previous congruent trial; Prev Incong: previous incongruent trial). Current trial congruency is indicated by line type (Incong: dashed line; Cong: black line). Reaction Time (in ms) is indicated on the y-axis. Error bars indicate ±1 standard error from the mean.

Critically, we observed a significant three-way interaction among modality condition, previous trial congruency, and current trial congruency, $F(1,31) = 5.45$, $p = 0.026$, $\eta_p^2 = 0.15$, $BF_{10} = 0.72$, $BF_{01} = 1.39$[10]. As in Experiment 1, and consistent with the orienting hypothesis, the CSE was larger in modality repeat trials (24 ms; $F(1,31) = 12.28$, $p = 0.001$, $\eta_p^2 = 0.28$, $BF_{10} = 2.60$, $BF_{01} = 0.38$) than in modality switch trials (2 ms; $F(1,31) < 1$, $BF_{10} = 0.26$, $BF_{01} = 3.91$) (Fig. 5). No other effects were significant (see Table 3).

---

[10] This interaction was also significant when considering only valid trials, $F(1,31) = 7.55$, $p = 0.01$, $\eta_p^2 = 0.20$, $BF_{10} = 0.94$, $BF_{01} = 1.07$, which is more comparable to Experiment 1 wherein all of the trials were valid.

Mean ER

We observed two significant effects. First, we observed a main effect of modality condition, $F(1,31) = 7.45$, $p = 0.010$, $\boxed{\eta_p^2} = 0.19$, $BF_{10} = 3.52$, $BF_{01} = 0.28$, because mean ER was higher in modality repeat trials (2.0%) than in modality switch trials (1.5%). Second, we observed a main effect of current trial congruency, $F(1,31) = 7.21$, $p = 0.012$, $\boxed{\eta_p^2} = 0.19$, $BF_{10} = 3.13$, $BF_{01} = 0.32$, because mean ER was higher in incongruent trials (2.1%) than in congruent trials (1.5%). Finally, we observed a non-significant trend toward a main effect of previous trial congruency, $F(1,31) = 2.92$, $p = 0.10$, $\boxed{\eta_p^2} = 0.097$, $BF_{10} = 0.38$, $BF_{01} = 2.62$, because mean ER was higher after congruent trials (1.9%) than after incongruent trials (1.6%). No other effects were significant (see Table 3).

Table 4-3. Statistics for repeated-measures and Bayesian ANOVAs in Exp. 3

Dependent	Effect	Repeated-measures			Bayesian	
		F	*p*	η_p^2	BF_{10}	BF_{01}
Mean RT	Modality Condition	11.22	.002	.27	3.07	0.33
	Previous Trial Congruency	3.88	.058	.11	0.26	3.85
	Current Trial Congruency	36.77	<.001	.54	1.84×10^{19}	5.44×10^{-20}
	Modality Condition x Previous Trial Congruency	2.13	.15	.06	0.24	4.22
	Modality Condition x Current Trial Congruency	1.91	.18	.06	0.25	4.01
	Previous Trial Congruency x Current Trial Congruency	4.84	.035	.14	0.64	1.57
	3-way interaction	5.45	.026	.15	0.72	1.39
	Modality Condition	7.45	.010	.19	3.52	0.28
Mean ER						
	Previous Trial Congruency	2.92	.097	.09	0.38	2.62
	Current Trial Congruency	7.21	.012	.19	3.13	0.32
	Modality Condition x Previous Trial Congruency	1.21	.28	.04	0.34	2.93
	Modality Condition x Current Trial Congruency	.02	.89	.00	0.18	5.45
	Previous Trial Congruency x Current Trial Congruency	2.46	.13	.07	0.62	1.60
	3-way interaction	.62	.44	.02	0.30	3.34

Discussion

In Experiment 3, we observed modality-specific CSEs coupled with a modality validity effect. Further, the modality validity effect was larger in Experiment 3 than in Experiment 2. These findings are consistent with the orienting hypothesis, which posits that orienting attention to a sensory modality can lead to the formation of a modality-specific task set. They also provide a potential mechanistic explanation, in terms of attentional orienting, for why the CSE was associated with modality-specific boundaries in Experiments 1 and 3 but not in Experiment 2.

However, it is worth mentioning that switching modalities reduced the CSE to a lesser degree when invalid trials were present than when they were not. Indeed, the effect size for the critical three-way interaction was much smaller in Experiment 3 (η_p^2 = 0.149) than in Experiment 1 (η_p^2 = 0.428). This suggests that even a small number of modality-invalid trials reduces the degree to which participants form modality-specific task sets. Future studies could explore whether this reduction varies parametrically with the percentages of modality-valid and modality-invalid trials.

General Discussion

We investigated whether context-defining features or task sets based on such features serve as boundaries for cognitive control processes underlying the CSE. Consistent with the latter hypothesis, Experiments 1 and 2 revealed that different sensory modalities (i.e., context-defining features) serve as boundaries for the CSE only when they are associated with different task sets. Further, Experiment 3 revealed a mechanism that may underlie the formation of a modality-specific task set: orienting attention to the sensory modality in which a distractor appears when participants expect a subsequent target in the same modality. To our knowledge, these findings are the first to distinguish between the attentional reset and task set hypotheses. They are also the first to explicitly investigate the relationship between attentional orienting and task set formation. Thus, they have important implications for current views of the control-driven CSE and for our understanding of how episodic memory interacts with control processes to engender this effect.

Implications for the task set and attentional reset hypotheses

The present findings provide novel support for the task set hypothesis (Egner, 2014; Hazeltine et al., 2011; Schumacher & Hazeltine, 2016; Spapé & Hommel, 2008). In this view, changing a context-defining feature reduces the CSE only when the updated feature is linked to a different task set (Hazeltine et al., 2011; Schumacher & Hazeltine, 2016). Consistent with this view, we observed modality-specific CSEs when the task structure allowed participants to categorize each trial using the sensory modality in which the task stimuli appeared and thereby form modality-specific task sets (i.e., in Experiments 1 and 3) (Schumacher & Hazeltine, 2016). Moreover, we did not observe modality-specific CSEs when the task structure did not allow participants to form modality-specific task sets via such a categorization (i.e., in Experiment 2). These findings support the view that task sets based on context-defining features (e.g., sensory modalities in the present study) serve as boundaries for cognitive control. In contrast, the present findings do not appear consistent with the attentional reset hypothesis (e.g., Kreutzfeldt et al., 2016). According to this view, changing the sensory modality in which the task stimuli appear signals that previous-trial control settings are no longer appropriate, which should always reduce or eliminate the CSE. Contrary to this view, however, we found that changing the sensory modality in which task stimuli appear does not reduce the CSE when participants cannot classify each trial as either "visual" or "auditory" (i.e., in Experiment 2). This result appears incompatible with the attentional reset hypothesis.

Given these findings, changing context-defining features other than the stimulus modality may also reduce the CSE by leading to a change in the task set that underlies performance. To test this hypothesis, researchers could conduct pairs of experiments analogous to the present Experiments 1 and 2. For example, imagine a pair of experiments based on a study conducted by Dignath et al. (2019), wherein the distractor and target in each trial were color words or color patches. As in that study, in the initial experiment the distractor and target in each trial would always appear in the same stimulus format (i.e., color word or color patch). Analogous to the present study, both the task set hypothesis and the attentional reset hypothesis would predict that a change in stimulus format across consecutive should reduce the CSE. In the second experiment, the distractor and target in each trial would appear equally often in the same format or in different formats (randomized across trials). Analogous to the present Experiment 2, this should prevent participants from classifying each trial as involving only "color words" or only "color patches" and, hence, prevent control processes from creating format-specific task sets. Thus, the task set hypothesis would predict no reduction of the CSE when the format changes across consecutive trials. In contrast, the attentional reset hypothesis would still predict such a reduction. Conducting such experiments would allow researchers to identify which features can serve as context-defining boundaries for the CSE and, critically, to determine whether or not the task set hypothesis can account for such boundaries.

Forming a modality-specific task set is associated with orienting attention

The findings from Experiment 3 reveal for the first time that forming a modality-specific task set is associated with orienting attention to the sensory modality associated with the distractor before the target appears. Specifically, we observed modality-specific CSEs coupled with a robust modality validity effect, consistent with orienting attention to the distractor's modality. Notably, we did not observe these effects in Experiment 2, wherein the distractor's modality did not systematically predict the target's modality. This suggests that control processes are more likely to employ context-defining features (e.g., different sensory modalities) to create distinct task sets when orienting to those features can systematically aid performance (e.g., by allowing participants to prepare for a target) than when it cannot.

This conclusion appears inconsistent with the idea that context-defining features are entirely task-irrelevant when they are independent of the stimulus-response mapping (Dignath et al., 2019). Instead, it suggests that even though context-defining features are irrelevant in terms of the task instructions, they become relevant to participants when they systematically predict the nature (e.g., the sensory modality) of upcoming task-relevant information. Consistent with this view, other findings also suggest that participants attend to irrelevant stimuli when they predict the nature of relevant information. For example, participants quickly learn contingencies between irrelevant and relevant stimuli when doing so allows them to predict the response to an upcoming target (Schmidt, Augustinova, & De Houwer, 2018; Schmidt & Besner, 2008). Future studies might therefore investigate the potential impact of attending to context-defining features on the presence and magnitude of context-specific CSEs.

Along these lines, functional neuroimaging data also suggest that attending to different context-defining features contributes to forming distinct task sets. Consider, for instance, a prior study wherein a cue in each trial specified contextual information that participants could use to identify an upcoming target (Cookson, Hazeltine, & Schumacher, 2016, 2019). Such cues engendered the largest performance benefit when participants employed a stimulus-response mapping that segregated different types of stimuli (i.e., faces and places) to different hands. Critically, functional magnetic resonance imaging data revealed increased functional connectivity between the stimulus- and response-related regions that were associated with a task set (e.g., sensory regions that process faces and motor regions that support movements of the right hand) when that task set was cued than when a different task set was cued. Consistent with the orienting hypothesis and with the present results, these findings suggest that participants oriented attention to the contextual information provided by the cues to form distinct task sets.

It is important to note, however, that some "context-specific" CSEs may not stem from orienting attention to a context-defining feature. For example, Spapé & Hommel (2008) reported that changing the gender (i.e., male or female) of the voice in which an auditory distractor word (i.e., "high" or "low") is spoken eliminates the CSE even when orienting to voice gender is unlikely to be helpful for identifying an upcoming target (i.e., a high- or low-pitched tone). Although their study contained feature integration confounds, this finding suggests that some changes in episodic context may reduce the CSE via mechanisms that differ from those posited by the orienting hypothesis. Future studies that employ confound-minimized task protocols could investigate this potentially important distinction.

Implications for current accounts of how control processes engender the CSE

The present findings show that forming a modality-specific task set is associated with orienting attention to the distractor's modality. However, the control-driven CSE that occurs within a task set appears to result from a modulation of the response cued by the distractor (Ridderinkhof, 2002; Stürmer, Leuthold, Soetens, Schröter, & Sommer, 2002; Weissman et al., 2014) rather than from a shift of attention toward the target (Botvinick, Braver, Barch, Carter, & Cohen, 2001; Braem, Verguts, & Notebaert, 2011; Gratton et al., 1992; Verguts & Notebaert, 2008, 2009). For example, when a long (i.e., 1000 ms) ISI separates the prime and probe in the visual-modality prime-probe task, which eliminates the overall congruency effect, the CSE is linked to a small positive congruency effect after congruent trials and a small negative congruency effect after incongruent trials (Weissman et al., 2015). Critically, the negative congruency effect favors a response modulation mechanism over an attentional shifting mechanism. For example, following incongruent trials inhibiting the response cued by the distractor could selectively delay responses in congruent trials wherein the target requires the inhibited response, resulting in a negative congruency effect. In contrast, even shifting all of one's attention toward the target and away from the distractor could eliminate, but not reverse, the congruency effect.

Given these considerations, the modality-specific CSEs that we observed in Experiments 1 and 3 may reflect modality-specific modulations of the response cued by the distractor. For example, an episodic memory of each trial may store not only whether a distractor was congruent or incongruent, but also whether it appeared in the visual or the auditory modality. Thus, when participants form modality-specific task sets, they may retrieve a memory of the previous trial that indicates both the modality and the congruency of the distractor (e.g., "the auditory distractor was congruent" or "the visual distractor was incongruent"). Such specificity could lead control processes to modulate the response cued by the current-trial distractor only if its modality matches the modality of the distractor in the previous trial. Future work could explore this idea.

Potential relevance of our findings to "conflict-specific" CSEs

The ability to store information about a distractor other than its congruency could also engender highly specific CSEs in other tasks. Consider hybrid Stroop-like tasks, for instance, wherein there are two types of distractors (e.g., a Stroop distractor and a Simon distractor) rather than just one (Egner, 2008). In such tasks, CSEs are often "distractor-specific", meaning that the congruency effect associated with each distractor is reduced following a trial in which that distractor was incongruent but not following a trial in which the other distractor was incongruent.

For example, in the hybrid Stroop-Simon task, the Stroop effect is reduced after a Stroop-incongruent trial but not after a Simon-incongruent trial while the opposite holds for the Simon effect (Egner, Delano, & Hirsch, 2007; Hommel, 1997; Jiang & Egner, 2014; Wendt, Kluwe, & Peters, 2006). A popular account of such distractor-specific CSEs is that they are "conflict-specific", meaning that different control processes are recruited to resolve different types of conflict (e.g., stimulus-stimulus Stroop conflict and stimulus-response Simon conflict) (Egner, 2008; Kornblum, Hasbroucq, & Osman, 1990)

The present results suggest another possible interpretation of distractor-specific CSEs. Namely, participants may store the contextual features that characterize each distractor (e.g., verbal features for a Stroop distractor and spatial features for a Simon distractor) along with the congruency of each distractor (i.e., congruent or incongruent) in their episodic memory of each trial (Weissman, in press). Consequently, retrieving an episodic memory of the previous trial may lead control processes to apply the control settings that were appropriate for each distractor in the previous trial to the distractor with the same contextual features in the current trial. Future studies that employ hybrid Stroop-like tasks could investigate this possibility.

A role for the local temporal context in cueing episodic memory retrieval

The present findings also bolster the view that the local temporal context can trigger the retrieval of an episodic memory of the previous trial (Egner, 2014). The local temporal context refers to the temporal frame or episode within which a participant perceives the current trial to occur. If this temporal frame or episode remains unchanged from the previous trial (e.g., because the participant is still employing the same task set), then it should cue the retrieval of a memory of the previous trial even in the absence of repeating stimulus, response, and/or context-defining features. The results of Experiment 2 provide novel support for this view by revealing a robust CSE even in the absence of stimulus, response, and sensory modality repetitions.

Limitations

One might argue that the present findings are specific to tasks wherein the distractor precedes the target. In line with this possibility, prior work has revealed that the CSE is larger when a distractor appears before a target, rather than with a target, consistent with the view that control processes need time to modulate the response cued by the distractor before the target response is executed (Weissman et al., 2015). However, even if the present findings are specific to protocols wherein the distractor appears before the target, the present tasks are still highly representative of those that engender robust, control-driven CSEs.

One might also argue that the present findings are specific to cross-modal paradigms. This seems unlikely because participants can orient attention not only to different sensory modalities but also to different contextual features within a single sensory modality (e.g., to spatial and non-spatial features) (Giesbrecht, Woldorff, Song, & Mangun, 2003). As we described earlier, however, future studies could investigate this possibility by determining whether our findings generalize to unimodal (e.g., visual modality) tasks.

Finally, one might argue that the presence of modality-specific CSEs in our experiments varies with the number of different trial types and/or their frequency of presentation, which may impact the overall complexity of the task. Prior findings, however, suggest no consistent relationship between CSE boundaries and task complexity. In some situations, CSE boundaries become larger when task complexity increases (e.g., see Experiment 3 vs. Experiment 2 of Hazeltine et al., 2011, which employed an 8-AFC task and a 4-AFC task, respectively). In others, they become smaller (e.g., see Experiment 4 vs. Experiment 1 from Hazeltine et al., 2011, which employed two stimulus categories vs. one stimulus category, respectively).

Critically, even if there is an influence of task complexity on the formation of CSE boundaries, our findings still support the task set hypothesis. Specifically, they show that task structure, which influences the formation of task sets (Hazeltine et al., 2011), is what determines the boundaries of control processes underlying the CSE (see also, Cookson, Hazeltine, and Schumacher, 2019). Interestingly, they also show that forming a CSE boundary is associated with orienting attention to the context-defining feature of the distractor. Investigating whether or not such orienting is critical for the formation of a CSE boundary could be an exciting topic for future research.

Conclusion

The present findings favor the task set hypothesis over the attentional reset hypothesis. Specifically, they indicate that CSE boundaries reflect task sets based on context-defining features, rather than context-defining features on their own. They also suggest that orienting attention to different context-defining features facilitates the formation of different task sets. Future studies that investigate how orienting attention to such features influences task set formation may therefore shed additional light on the boundaries of cognitive control.

Chapter 5 Rethinking Attentional Reset: Task Sets Determine the Boundaries of Adaptive Control

Abstract

Adaptive control processes that minimize distraction often operate in a context-specific manner. For example, they may minimize distraction from irrelevant conversations during a lecture but not in the hallway afterwards. It remains unclear, however, whether (a) salient perceptual features or (b) task sets based on such features serve as contextual boundaries for adaptive control in standard distractor-interference tasks. To distinguish between these possibilities, we manipulated whether the structure of a standard, visual distractor-interference task allowed (Experiment 1) or did not allow (Experiment 2) participants to associate salient visual features (i.e., color patches and color words) with different task sets. We found that changing salient visual features across consecutive trials reduced a popular measure of adaptive control in distractor-interference tasks – the congruency sequence effect (CSE) – only when the task structure allowed participants to associate these visual features with different task sets. These findings extend prior support for the task set hypothesis from somewhat atypical cross-modal tasks to a standard unimodal task. In contrast, they pose a challenge to an alternative "attentional reset" hypothesis, and related views, wherein changing salient perceptual features always results in a contextual boundary for the CSE.

Introduction

Humans match new information to prior experiences in order to predict future events. While drinking a cup of coffee at a local café, for example, a student may recall an associated memory of a friend who usually arrives at the café at around the same time. The student may then prepare to purchase a second cup of coffee expecting the friend to arrive soon. On some occasions, however, a change in episodic context requires humans to override previous control settings. For instance, if the student's friend has another appointment one morning, then the student should resist the urge to purchase an extra cup of coffee. As this example shows, the ability to flexibly adapt to changing episodic contexts is an important aspect of adaptive control, which is crucial for enabling goal-driven behavior (Miller & Cohen, 2001).

To investigate adaptive control in the laboratory, researchers use distractor-interference tasks (Eriksen & Eriksen, 1974; Eriksen & Schultz, 1979; Kunde & Wuhr, 2006; Simon & Rudell, 1967). In each trial of the prime-probe task, for example, a distractor precedes a target that study participants are supposed to identify. Participants respond more slowly in incongruent trials, wherein the distractor and target indicate different responses, than in congruent trials, wherein these stimuli indicate the same response. However, this congruency effect is smaller after incongruent trials than after congruent trials. This *congruency sequence effect* (CSE) indicates that participants are less distractible just after they experience heightened distraction in an incongruent (vs. congruent) trial (Gratton, Coles, & Donchin, 1992; Botvinick, Braver, Barch, Carter, & Cohen, 2001; Egner, 2008).

However, other processes may also contribute to the CSE under certain conditions. First, feature integration processes may engender a CSE when the nature of stimulus and/or response repetitions across trials is confounded with trial congruency (Hommel et al., 2004; Mayr, Awh,

& Laurey, 2003). For example, in 2-alternative-forced-choice (2-AFC) tasks, exact stimulus and response repetitions occur when trial congruency repeats in consecutive trials (e.g., red-red -> red-red) while partial repetitions occur when trial congruency alternates (e.g., red-red -> red-blue). Since response times are faster for exact repetitions than for partial repetitions, this confound engenders a CSE.

Second, contingency learning processes may engender a CSE when a distractor's identity predicts a congruent response to the target with greater-than-chance accuracy (Schmidt & DeHouwer, 2011). This occurs in tasks with more than two possible stimuli and responses wherein there are fewer unique congruent stimuli than unique incongruent stimuli. To equate the number of congruent and incongruent trials, researchers sometimes present each unique congruent stimulus more frequently than each unique incongruent stimulus (e.g., red-red might appear three times as often as red-blue). This procedure leads participants to prepare the "high-contingency" congruent response more strongly than each possible "low-contingency" incongruent response, which increases the size of the congruency effect. Critically, such preparation is greater when the previous trial was congruent relative to incongruent. Thus, contingency learning processes can engender a CSE.

Critically, researchers can employ confound-minimized paradigms to observe a CSE even in the absence of such learning and memory processes (Jimenez & Mendez, 2014; Kim & Cho, 2014; Schmidt & Weissman, 2014). Such protocols allow researchers to isolate the contribution of cognitive control processes to the CSE.

The episodic retrieval view of the CSE

The *episodic retrieval view* posits that a CSE occurs when the current trial triggers the retrieval of a memory of the previous trial (Dignath, Johannsen, Hommel, & Kiesel, 2019; Egner, 2014; Frings et al., 2020; Spapé & Hommel, 2008; Weissman, Hawks, & Egner, 2016; Hazeltine, Lightman, Schwarb, & Schumacher, 2011). In this view, various features (i.e., stimuli, responses, salient perceptual features, task sets, etc.) are "bound" within an episodic memory of the previous trial (Hommel, 2004). Reencountering any of these features in the current trial then triggers the retrieval of the remaining features such as, for example, the control settings that a participant employed to minimize distraction in a previous incongruent trial. This biases control processes to adopt the same settings (e.g., to inhibit the response cued by the distractor), leading to a smaller congruency effect after incongruent trials than after congruent trials (i.e., a CSE).

The episodic retrieval view predicts that changing salient perceptual features across consecutive trials should reduce the CSE by impairing the ability to retrieve an episodic memory of the previous trial (Spapé & Hommel, 2008). In line with this prediction, changing the stimulus format (i.e., color words vs. color patches; Dignath et al., 2019), the voice gender of a spoken, auditory distractor word (i.e., male vs. female; Spapé & Hommel, 2008), the stimulus color (Braem, Hickey, Duthoo, & Notebaert, 2014), or even the entire task (Kiesel, Kunde, & Hoffmann, 2006) reduces the CSE. As we describe next, however, two variants of the episodic retrieval view posit that such contextual boundaries for the CSE arise for different reasons.

Before proceeding, however, it is important to define three terms. First, we define a "task" as a set of cognitive operations and stimulus-response mappings that a participant can implement to respond quickly and accurately to goal-relevant stimuli (Rogers & Monsell, 1995). For instance, in each trial of a Stroop color word task, participants employ a set of cognitive operations and stimulus-response mappings to identify the ink color (e.g., red) in which an irrelevant color word (e.g., GREEN) appears. Second, we define a “task set” as an intention to perform a task, which involves maintaining a task in working memory (Rogers & Monsell, 1995). For instance, to perform the Stroop task, participants need to retrieve a representation of the task from long-term memory and keep it active in working memory during the experiment. Third, we define a “dimension” as a feature of a stimulus (e.g., its color, location, sensory modality, etc.) that can take on any of multiple values (e.g., red, blue, or green for the color dimension) (Kornblum et al., 1990; Egner, 2008). For instance, in the Stroop task the task-relevant stimulus dimension is ink color and the task-irrelevant stimulus dimension is word identity.

The attentional reset hypothesis

According to the *attentional reset hypothesis*, changing salient perceptual features associated with the task-relevant stimulus dimension (i.e., target-related information) reduces the CSE by triggering the formation of a new event, or episode, in working memory (Kreutzfeldt, Stephan, Willmes, & Koch, 2016). Here, changing salient perceptual features renders previous-trial control processes irrelevant, leading to an attentional reset in the current trial and, consequently, a reduced CSE. Critically, a change in perceptual features should trigger an attentional reset even when the overall task remains the same across consecutive trials. Thus, the attentional reset hypothesis posits that salient perceptual features on their own serve as boundaries for the CSE.

The attentional reset hypothesis fits with views, developed largely using unimodal visual tasks, wherein adaptive control processes underlying the CSE are specific to the task-relevant stimulus dimension. For instance, consider the influential conflict-monitoring view. Here, a CSE occurs because control processes increase attention to the task-relevant stimulus dimension after high-conflict incongruent trials (Botvinick et al., 2001). This reduces the influence of the distractor on performance in the next trial and thereby minimizes the congruency effect. Critically, according to Braem, Abrahamse, Duthoo, & Notebaert (2014), the conflict-monitoring model implies that a CSE should not appear if the task-relevant stimulus dimension – a highly salient perceptual feature – changes (for a related view, see Egner, 2014). These authors argue that since the heightened focus of attention after incongruent trials is specific to that dimension, it should vanish if that dimension changes. In line with this possibility, changing the task-relevant stimulus dimension can eliminate indices of adaptive control including the CSE (Notebaert & Verguts, 2008; Wühr, Duthoo, & Notebaert, 2015). However, the presence of feature integration confounds in these prior studies makes it difficult to draw firm conclusions about whether such effects index contextual boundaries for (a) adaptive control or (b) experimental confounds.

The task set hypothesis

The *task set hypothesis* posits that changing salient perceptual features reduces the CSE only when it leads participants to change the task set (i.e., representation of the current task in working memory) that underlies performance (Grant, Cookson, & Weissman, 2020; Hazeltine et al., 2011; Schumacher & Hazeltine, 2016). In this view, participants use salient perceptual features (e.g., the sensory modality in which task stimuli appear) to distinguish between different subsets of trials. This allows participants to divide a complex task (e.g., a cross-modal task with visual and auditory trials) into two or more simpler tasks (e.g., a visual task and an auditory task), which reduces the complexity of the overall stimulus-response (S-R) mapping (Hazeltine et al., 2011). Critically, switching to a new task set disrupts the ability to retrieve an episodic memory of the previous trial, thereby reducing the CSE (Hazeltine et al., 2011; Spapé & Hommel, 2008). In other words, task sets based on salient perceptual features, rather than such features on their own, serve as boundaries for the CSE.

Recent findings from a confound-minimized, cross-modal prime-probe task that we designed support the task set hypothesis (Grant et al., 2020). Similar to Hazeltine et al. (2011, Experiment 1), in our first experiment the initial distractor and subsequent target both appeared in (a) the visual modality (50% of trials) or (b) the auditory modality (50% of trials). Here, participants could easily classify each trial as “visual” or “auditory”. They could also orient to the sensory modality in which the distractor appeared to prepare for the upcoming target, which always appeared in the same modality. Therefore, they were likely to form modality-specific task sets. In our second experiment, the sensory modality of the distractor varied independently of the sensory modality of the target. Here, participants could not (a) classify each trial as visual or auditory, or (b) orient to the distractor’s modality to prepare for the upcoming target. Consequently, they were likely to form a single, modality-general task set. Consistent with the task set hypothesis, but not with the attentional reset hypothesis, we found that a change in sensory modality reduced the CSE in Experiment 1 but not in Experiment 2.

Does an attentional reset explain contextual boundaries for the CSE in unimodal tasks?

Despite such evidence supporting the task set hypothesis, some findings suggest that the attentional reset hypothesis may explain contextual boundaries for the CSE in unimodal tasks.

These findings show that changing salient perceptual features in unimodal tasks reduces the CSE even when orienting to the perceptual features of the distractor – the basis of task set formation in our cross-modal study (Grant et al., 2020) – cannot facilitate identification of the target (Spapé & Hommel, 2008; Braem, Hickey, et al., 2014). For instance, consider a prior study wherein participants indicated whether an auditory target was a high- or low-frequency tone while ignoring an irrelevant auditory distractor word ("high" vs. "low") spoken in a male or female voice (Spapé & Hommel, 2008). Although orienting to the voice gender of the auditory distractor cannot aid identification of the target, changing the voice gender in consecutive trials reduces the CSE. Analogous findings from a unimodal visual task indicate that changing the stimulus color in consecutive trials reduces the CSE even when this feature is completely irrelevant to identifying a relevant target (Braem, Hickey, et al., 2014). These findings suggest the possibility that, in unimodal tasks, salient perceptual features on their own – rather than task sets based on such features – serve as contextual boundaries for the CSE. However, the presence of feature integration confounds in these studies makes it difficult to determine whether salient perceptual features serve as boundaries for (a) adaptive control or (b) experimental confounds.

The present study

The goal of the present study is to distinguish between the task set and attentional reset hypotheses in a standard visual task that lacks both feature integration and contingency learning confounds. To achieve this goal, we employ a prime-probe task wherein the stimuli are color words and color patches (Dignath, Johannsen, Hommel, & Kiesel, 2019). In Experiment 1, the task structure allows participants to associate these distinct stimulus formats (i.e., salient perceptual features) with different task sets. Thus, both hypotheses predict a smaller CSE when the stimulus format changes than when it repeats as in Dignath et al. (2019). In Experiment 2, the task structure does not allow participants to associate these stimulus formats with different task sets. Thus, while the attentional reset hypothesis continues to predict a smaller CSE when the stimulus format changes (vs. repeats), the task set hypothesis no longer predicts such a reduction.

Experiment 1

In Experiment 1, we conducted a conceptual replication of Experiment 2 from Dignath et al. (2019). As we described earlier, in this experiment both hypotheses predict a smaller CSE when the stimulus format changes than when it repeats. Therefore, we reasoned that replicating this finding, which has only been reported in one prior study, would provide important information (e.g., about effect sizes) for subsequently distinguishing between the task set and attentional reset hypotheses in a second experiment.

Methods

Participants

Prior to data collection, we registered all hypotheses, methods, and analyses on the Open Science Framework (OSF) (https://osf.io/8p5n6). As described there, we conducted a power analysis (G*Power 3.1.9.4; Faul, Erdfelder, Lang, & Buchner, 2007) based on the three-way interaction among format transition, current trial congruency, and previous trial congruency (α = 0.05, η_p^2 = 0.29) that Dignath et al. (2019) observed in their second experiment. The results of this analysis indicated that collecting usable data from 48 participants would provide more than 95% power to observe the same interaction in the present experiment. The University of Michigan's Behavioral Sciences Internal Review Board (IRB) approved the study protocol.

Fifty-three undergraduates from the University of Michigan participated for course credit. We excluded data from a total of five participants. Four self-reported neurological disorders and/or using psychoactive medications and one performed the task with less than 75% accuracy. None of the remaining 48 participants (11 male, 37 female; 44 right-handed, 3 left-handed; 1 ambidextrous; age range: 18–23 years; mean age: 18.75 years, standard deviation of age: 1 year) reported any history of head trauma, uncorrected visual or hearing impairments, seizures, or neurological disorders. We uploaded the raw data from these 48 participants to the OSF (https://osf.io/354tj/?view_only=76564e5f17a0419bb9eb9fb58d2a7e0b).

Stimuli and Apparatus

The stimuli were color words and color patches. As in Dignath et al. (2019), distractor color words (Red: 3.1° x 1.6°, Blue: 3.6° x 1.6°, Green: 4.7° x 1.6°, Yellow: 5.2° x 1.6°) and distractor color patches (RED: 1.6° x 1.6°, BLUE: 1.6° x 1.6°, GREEN: 1.6° x 1.6°, YELLOW: 1.6° x 1.6°) were larger than target color words (Red: 2.1° x 1.0°, Blue: 2.6° x 1.0°, Green: 3.1° x 1.0°, Yellow: 3.6° x 1.0°) and target color patches (RED: 1.0° x 1.0°, BLUE: 1.0° x 1.0°, GREEN: 1.0° x 1.0°, YELLOW: 1.0° x 1.0°), respectively. We employed the Psychophysics Toolbox (Brainard, 1997), a MATLAB plug-in, on three Windows PC computers to present the task stimuli and to record participants' responses. We positioned each study participants' eyes approximately 55 cm away from the computer screen.

Experimental Design

The experiment consisted of a 64-trial practice block followed by ten test blocks with 96 trials each. Each block started and ended with a 2 s fixation cross (0.48° x 0.48°). In each 2 s trial (Fig.1), there were four sequential events: (1) a distractor (duration, 133 ms) (2) a blank screen (duration, 33 ms), (3) a target (duration, 133 ms), and (4) a second blank screen (duration, 1700 ms). In half the trials, the distractor and target were color words (red, blue, green, or yellow) while in the other half they were color patches (red, blue, green, or yellow). All of the stimuli appeared on a black background.

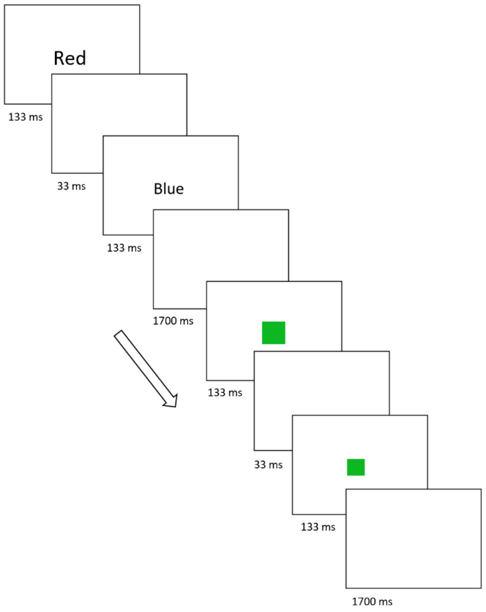


Figure 5-1. The prime-probe task in Experiment 1

This figure illustrates two sequential trials: an incongruent trial with color word stimuli followed by a congruent trial with color patch stimuli. The number beneath each box indicates the length of the corresponding trial component in milliseconds (ms). The arrow indicates the passage of time.

We created sixteen distractor-target pairs (i.e., trial types). Eight consisted of color patches and eight consisted of color words. Within each set of eight pairs, four were congruent (red-red; blue-blue; green-green; yellow-yellow) and four were incongruent (red-blue; blue-red, green-yellow; yellow-green). Each of these pairs preceded and followed each other equally often, separately for odd and even trials. In each block, one pair appeared one less time than the other pairs since no trial preceded the first trial. However, since we created a unique trial sequence for every block, the less-frequent pair varied randomly across blocks.

To isolate adaptive control processes, we employed a confound-minimized task protocol (Schmidt & Weissman, 2014; Weissman, Egner, Hawks, & Link, 2015; Weissman, Jiang, & Egner, 2014). This involved creating a pair of 2-alternative-forced-choice (2-AFC) tasks: one with red and blue words and patches and one with green and yellow words and patches. To avoid feature integration (e.g., stimulus repetition) confounds (Hommel et al., 2004; Mayr et al., 2003), we presented stimuli from the red-blue task in the odd trials of each block and stimuli from the green-yellow task in even trials of each block. To avoid contingency learning (i.e., stimulus frequency) confounds, we presented the two congruent and two incongruent distractor-target pairs in each 2-AFC task equally often in every block (Schmidt & De Houwer, 2011).

Procedure

Participants provided informed written consent prior to starting the experiment. Next, a research assistant read the task instructions to the participant. The instructions indicated that participants should respond as quickly and as accurately as possible to the second stimulus (i.e., the target) in each trial while ignoring the first stimulus (i.e., the distractor). In particular, participants were instructed to respond to a red, blue, green, or yellow patch/word target by pressing the “D” (right index finger), “F” (right middle finger), “G” (right ring finger), or “H” (right pinkie finger) key, respectively, on a QWERTY keyboard. If a participant pressed the wrong key in a given trial, or failed to respond within 1500 ms, the word “Error” appeared centrally for 200 ms.

Data Analyses

In our analyses of mean reaction time (RT), we excluded (1) practice trials, (2) the first trial of each block, (3) outliers in RT greater than $3S_n$ (Rousseeuw & Croux, 1993)[11] from the

---

[11] We employed a robust estimator of scale called S_n to identify outliers in RT because of its relatively low type I and type II error rates relative to other estimators of scale (e.g., standard deviation; Jones, 2019).

conditional mean, (4) trials with omitted or incorrect responses (errors), and (5) trials immediately following errors. We excluded the same trial types from our analyses of mean error rate (ER) with the exception of errors, which were the dependent measure. On average, 9.9% of the trials were errors and 4.3% were outliers.

Following these exclusions, we calculated mean RT and mean ER for each trial type. We then conducted separate repeated-measures ANOVAs on mean RT and mean ER. In each ANOVA there were four within-participants factors: current trial stimulus format (color patch, color word), format transition (repeat, switch), previous trial congruency (congruent, incongruent), and current trial congruency (congruent, incongruent). We included current trial stimulus format as a factor only to account for the extra variance that it produced and, hence, do not report findings related to this factor. We note, however, that this factor did not influence the critical three-way interaction among format transition, previous trial congruency, and current trial congruency.

Results

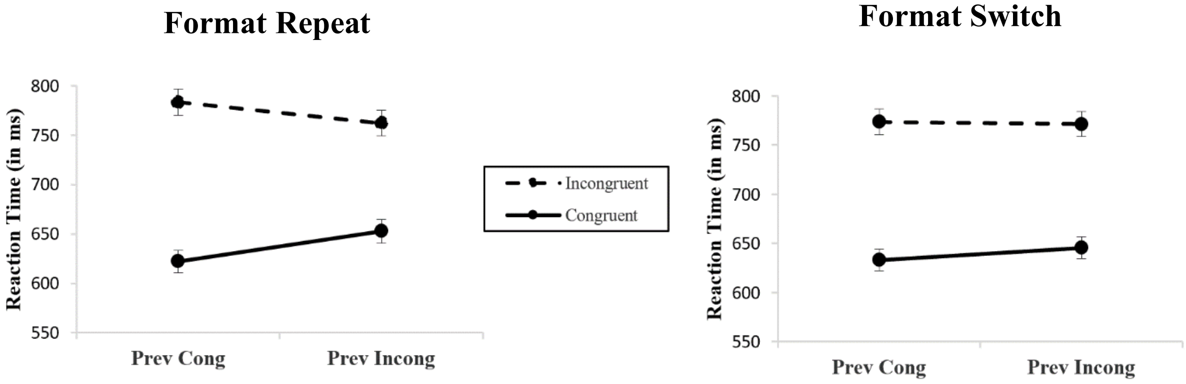
Mean RT

There were three significant effects. First, we observed a main effect of current trial congruency (i.e., a congruency effect), $F(1,47) = 345.17$, $p < 0.001$, $\eta_p^2 = 0.88$, because mean RT was longer in incongruent trials (774 ms) than in congruent trials (639 ms). Second, we observed a two-way interaction between previous trial congruency and current trial congruency (i.e., a CSE), $F(1,47) = 37.72$, $p < 0.001$, $\eta_p^2 = 0.45$, because the congruency effect was smaller after incongruent trials (120 ms) than after congruent trials (151 ms). Third, we observed an interaction among format transition, previous trial congruency, and current trial congruency,

$F(1,47) = 8.41$, $p = 0.006$, $\eta_p^2 = 0.15$. As expected, the CSE was larger when the stimulus format repeated across consecutive trials (47 ms; $F(1,47) = 32.67$, $p < 0.001$, $\eta_p^2 = 0.41$) than when the stimulus format switched (15 ms; $F(1,47) = 5.03$, $p = 0.030$, $\eta_p^2 = 0.10$) (Fig. 2).

Mean ER

We observed a significant main effect of current trial congruency, $F(1,47) = 16.83$, $p < 0.001$, $\eta_p^2 = 0.26$. As expected, mean ER was higher in incongruent relative to congruent trials (8.65% vs. 5.99%). No other effects were significant.



Format Repeat

Format Switch

Figure 5-2. CSE in each of the two critical trial types of Experiment 1

Previous trial congruency varies on the x-axis (Prev Cong: previous congruent trial; Prev Incong: previous incongruent trial). Current trial congruency varies with line type (dashed line: incongruent; black line: congruent). Reaction time (in ms) appears on the y-axis. Error bars indicate ± 1 standard error of the mean.

Exploratory Analysis

Frings and colleagues (2020) proposed that the influence of episodic memory on adaptive control reflects two processes: (1) *binding*, or encoding, distinct features within an episodic memory of the previous trial and (2) *retrieving* this memory in the current trial. Therefore, one may wonder whether episodic binding in the previous trial and episodic retrieval in the current trial separately influence the formation of a CSE boundary. There are two logical possibilities. First, episodic binding in the previous trial may be stronger for some perceptual features (e.g., color words) than for others (e.g., color patches). This view predicts a *larger* reduction of the CSE when the stimulus format changes if the *previous trial* contains a strongly (vs. weakly) bound stimulus format. That is, greater episodic binding for one stimulus format than for another in the previous trial should lead to an interaction among *previous-trial stimulus format*, format transition, previous trial congruency, and current trial congruency. A second possibility is that some perceptual features (e.g., color words) serve as stronger retrieval cues for an episodic memory of the previous trial than others (e.g., color patches). This view predicts a *smaller* reduction of the CSE when the stimulus format changes if the *current trial* contains a stimulus format that serves as a more (vs. less) effective retrieval cue. That is, greater episodic retrieval triggered by one stimulus format than by another in the current trial should lead to an interaction among *current-trial stimulus format*, format transition, previous trial congruency, and current trial congruency.

To explore these possibilities, we conducted two within-participants ANOVAs on mean RT. The factors in the first ANOVA were *previous trial stimulus format* (color patch, color word), format transition (repeat, switch), previous trial congruency (congruent, incongruent), and current trial congruency (congruent, incongruent). The factors in the second ANOVA were *current trial*

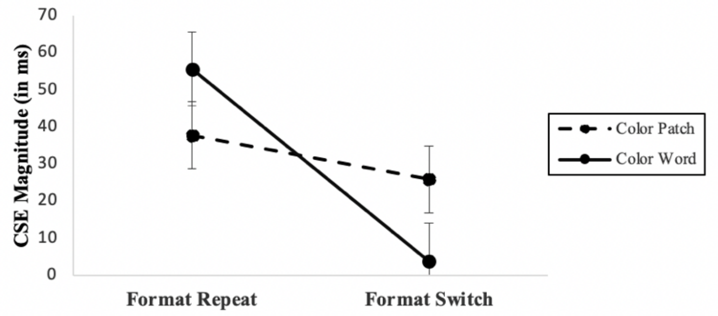


Figure 5-3. The CSE by format transition and previous trial stimulus format in Exp. 1

Format transition varies on the x-axis. Previous trial format varies by line type. CSE Magnitude (in ms) appears on the y-axis. Error bars indicate ±1 standard error of the mean.

format (color patch, color word), format transition (repeat, switch), previous trial congruency (congruent, incongruent), and current trial congruency.

The first ANOVA revealed a significant four-way interaction among *previous trial stimulus format*, format transition, previous trial congruency, current trial congruency, $F(1,47) = 5.22$, $p = 0.027$, $\eta_p^2 = 0.10$. We observed this interaction because the difference in CSE magnitude was larger after color-word trials (52 ms; $F(1,47) = 12.79$, $p < 0.001$, $\eta_p^2 = 0.21$) than after color-patch trials (12 ms, $F < 1$) (see Fig. 3). Follow-up analyses of the data from format-switch trials revealed that the CSE remained significant after color-patch trials (26 ms, $F(1,47) = 8.16$, $p = 0.006$, $\eta_p^2 = 0.15$) but not after color-word trials (4 ms, $F < 1$). These findings suggest that differences in *episodic binding* related to the two stimulus formats influenced the formation of CSE boundaries.

The second ANOVA did not reveal a significant four-way interaction among *current trial stimulus format*, format transition, previous trial congruency, current trial congruency ($F < 1$). Thus, there was no evidence to suggest that differences in *episodic retrieval* related to the two stimulus formats influenced the formation of CSE boundaries. Given that these were exploratory analyses, however, future studies should investigate the dissociation we observed in an *a priori* fashion.

Discussion

In line with both hypotheses, changing the stimulus format in Experiment 1 reduced the CSE. This finding replicates the key result of Dignath et al. (2019). Exploratory analyses further revealed that changing the stimulus format reduced the CSE more when color words (vs. color patches) appeared in the previous trial. However, it remains unclear whether task sets based on stimulus formats or stimulus formats on their own serve as boundaries for the CSE. We distinguish between these possibilities in Experiment 2.

Experiment 2

The goal of Experiment 2 was to distinguish between the task set and attentional reset hypotheses. To this end, we independently varied the stimulus format (color patch or color word) of (a) the distractor and (b) the target. Consequently, the distractor and target in each trial were equally likely to appear in the same format (i.e., color patch/color patch or color word/color word) or in different formats (i.e., color patch/color word or color word/color patch). Under such conditions, participants cannot classify each trial as involving only color words or only color patches. They also cannot orient to the format of the distractor to prepare for a target that always appears in the same format. Therefore, they are unlikely to create format-specific task sets.

Critically, the task set and attentional reset hypotheses make opposing predictions. The task set hypothesis predicts that changing the stimulus format will not reduce the CSE, because the task structure prevents participants from creating format-specific task sets. In other words, participants should form a single, format-general task set and, consequently, a robust CSE should appear regardless of whether the stimulus format changes or repeats. In contrast, the attentional reset hypothesis predicts that changing the target's stimulus format will reduce the CSE exactly as in Experiment 1. According to this hypothesis, changing the target's stimulus format biases the system to create a new episode, which signals that previous-trial control settings are no longer relevant and, consequently, "resets" attentional processes (Kreuzfeldt et al., 2016).

Methods

Participants

As in Experiment 1, we registered all hypotheses, methods, and analyses on the OSF (https://osf.io/mz24a) prior to data collection. To determine an appropriate sample size, we conducted a power analysis (G*Power 3.1.9.4; Faul et al., 2007) using the three-way interaction among format transition, current trial congruency, and previous trial congruency in Experiment 1 (α = 0.05, η_p^2 = 0.15). The results indicated that collecting usable data from 70 participants would provide over 90% power to observe an analogous three-way interaction among target format transition, previous trial congruency, and current trial congruency in the present experiment[12]. The University of Michigan's Behavioral Sciences IRB approved the study protocol.

Eighty-one undergraduates from the University of Michigan participated for course credit. We excluded the data from 11 participants due to computer issues (two participants) and for performing the task with less than 75% accuracy (nine participants). None of the remaining 70 participants (21 male, 49 female; 68 right-handed, 2 left-handed; age range: 18–21 years; mean age: 18.74 years, standard deviation of age: 0.88 years) reported any history of head trauma, uncorrected visual or hearing impairments, seizures, or neurological disorders. We uploaded the raw data from these 70 participants to the OSF (https://osf.io/jx74f/?view_only=f3fd29082f23446bb21fa6e3ddfa7432).

Stimuli and Apparatus

---

[12] In our preregistration, we proposed to examine only the four-way interaction. However, examining the three-way interaction provides a more specific test of the view that adaptive control processes are specific to the task-relevant stimulus dimension (Kreutzfeldt et al., 2016). It is also more consistent with our use of the three-way interaction from Experiment 1 to guide the power analysis.

The stimuli and apparatus were identical to those in Experiment 1.

Experimental Design

The experimental design was the same as that in Experiment 1 with two exceptions. First, to vary the stimulus format of (a) the distractor (color patch, color word) and (b) the target (color patch, color word) *independently* across trials, we created 32 (rather than 16) distractor-target pairs. Second, in each block we employed a first-order counterbalanced trial sequence in which there were four “same format” trial types (word congruent, word incongruent, patch congruent, patch incongruent) and four “different format” trial types (word-patch congruent, word-patch incongruent, patch-word congruent, and patch-word incongruent). These trial types preceded and followed each other equally often in every block – separately in odd and even trials - except for one that appeared one time less because no trial preceded the first trial of each block. Finally, as in Experiment 1, we created a different trial sequence for every block, separately for each participant.

Procedure

The procedure was the same as in Experiment 1 with one exception. The experiment consisted of a 64-trial practice block followed by five 128-trial (versus ten 64-trial) test blocks. We inserted a rest break halfway through each block to equate the number and frequency of rest breaks to those in Experiment 1.

Data Analyses

We excluded the same trials from our analyses as in Experiment 1. On average, 11.4% of the trials were errors and 2.6% were outliers. Following these exclusions, we calculated mean RT and mean ER for each trial type. We then conducted separate repeated-measures ANOVAs on mean RT and mean ER. In each ANOVA, there were four within-participants factors: distractor format transition (repeat, switch), target format transition (repeat, switch), previous trial congruency (congruent, incongruent) and current trial congruency (congruent, incongruent).

Results

Mean RT

We observed two significant main effects. First, we observed a main effect of previous trial congruency, $F(1, 69) = 7.70$, $p = 0.007$, $\eta_p^2 = 0.10$, because mean RT was longer after incongruent trials (661 ms) than after congruent trials (657 ms). This "post-conflict" slowing is commonly observed in distractor-interference tasks and likely reflects slower response times following conflict to reduce the possibility of making an erroneous response (Verguts, Notebaert, Kunde, & Wühr, 2011; Ullsperger et al., 2005). Second, we observed a main effect of current trial congruency, $F(1,69) = 419.31$, $p < 0.001$, $\eta_p^2 = 0.86$, because mean RT was longer in incongruent trials (704 ms) than in congruent trials (613 ms).

We also observed a pair of significant interactions. First, we observed an interaction between target format transition and previous trial congruency, $F(1, 69) = 8.30$, $p = 0.005$, $\eta_p^2 = 0.11$. Although mean RT was always longer after incongruent (vs. congruent), this effect was larger when the target format switched (7 ms; $F(1,69) = 14.6$, $p < 0.001$, $\eta_p^2 = 0.17$) than when it repeated (1 ms; $F < 1$). Second, we observed an interaction between previous trial congruency and current trial congruency (i.e., a CSE), $F(1, 69) = 80.60$, $p < 0.001$, $\eta_p^2 = 0.54$. As expected, the congruency effect was smaller after incongruent trials (78 ms; $F(1,69) = 320.0$, $p < 0.001$, $\eta_p^2 = 0.82$) than after congruent trials (103 ms; $F(1,69) = 436.0$, $p < 0.001$, $\eta_p^2 = 0.86$).

No other effects were significant including the three-way interactions among distractor format transition, previous trial congruency, and current trial congruency ($F < 1$) and target format transition, previous trial congruency, and current trial congruency ($F < 1$). The four-way interaction was also not significant, $F(1, 69) = 1.82$, $p = 0.18$, $\eta_p^2 = 0.026$)[13]. Indeed, the CSE was similarly robust when both formats repeated (30 ms), only the distractor's format repeated (22 ms), only the target's format repeated (24 ms), and when neither the distractor's format nor the target's format repeated (28 ms) (all four p values < 0.001; see Fig. 4).

---

[13] Exploratory analyses revealed that this four-way interaction remained non-significant ($F < 1$) when we included only the same trial types as in Experiment 1, wherein the distractor and target always appeared in the same format.

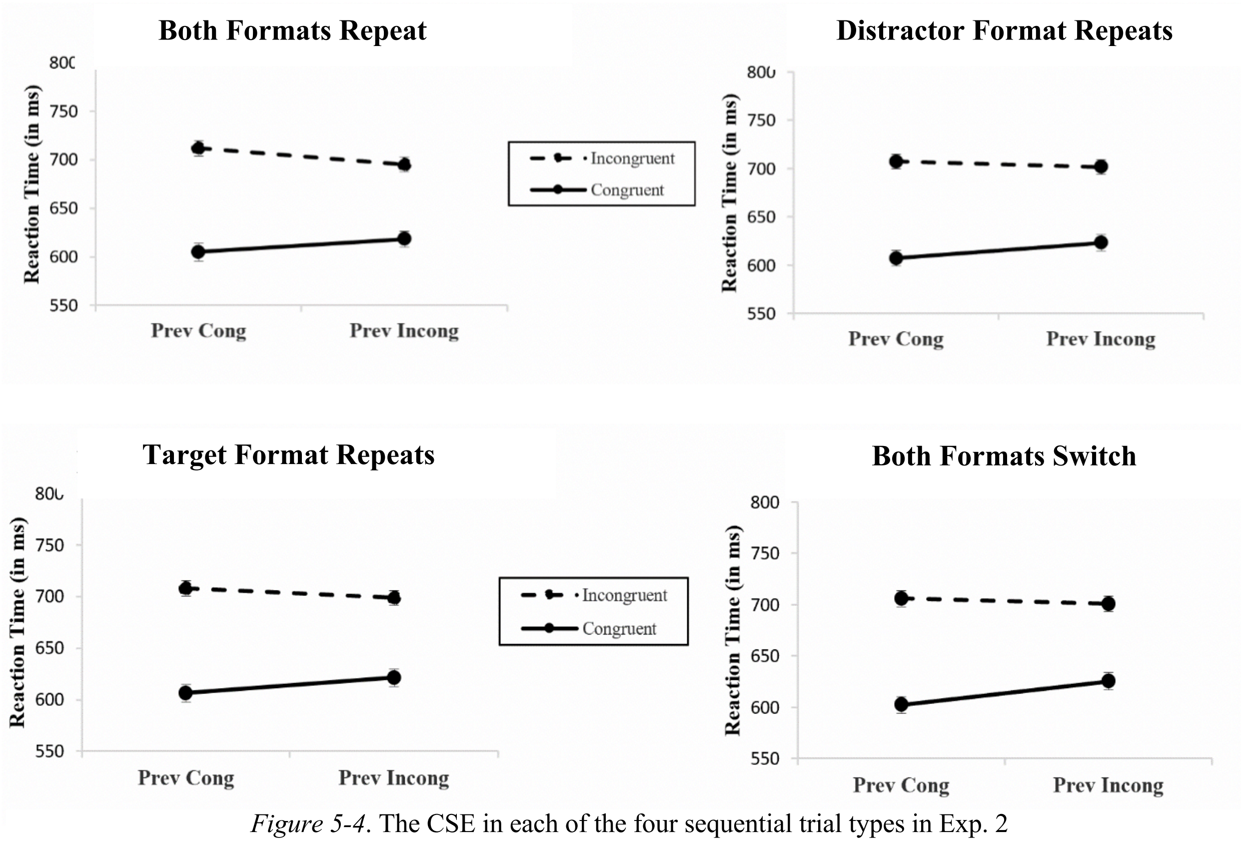


Figure 5-4. The CSE in each of the four sequential trial types in Exp. 2

Previous trial congruency varies on the x-axis (Prev Cong: previous congruent trial; Prev Incong: previous incongruent trial). Current trial congruency varies by line type (Dashed line: incongruent; Black line: congruent). Reaction time (in ms) appears on the y-axis. Error bars indicate ± 1 standard error of the mean.

Mean ER

We observed two significant main effects. First, we observed a main effect of previous trial congruency, $F(1, 69) = 8.21$, $p = 0.006$, $\eta_p^2 = 0.11$, because mean ER (a) was lower after incongruent trials (8.0%) than after congruent trials (8.7%). Second, we observed a main effect of current trial congruency, $F(1, 69) = 37.16$, $p < 0.001$, $\eta_p^2 = 0.35$, because mean ER was lower in congruent trials (7.0%) than in incongruent trials (9.7%). No other effects were significant.

Exploratory Across-Experiment Comparison

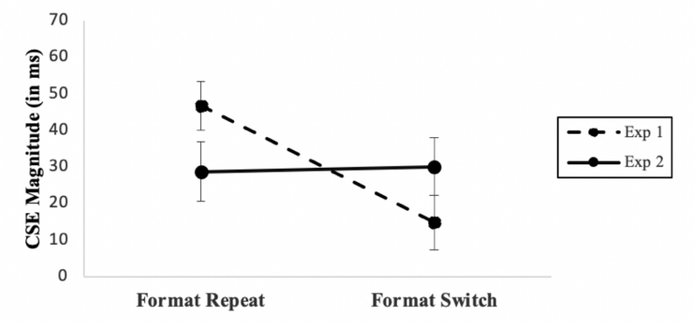


Figure 5-5. Across-experiment analysis

Whether the format of the distractor and target both repeated or both switched varies on the x-axis, separately for Experiment 1 (dotted line) and Experiment 2 (dashed line). The y-axis plots CSE magnitude (in ms) in each of these four conditions.

Our results thus far support the task set hypothesis: changing the stimulus format reduces the CSE only when the task structure allows participants to create format-specific task sets (i.e., in Experiment 1 but not Experiment 2). As a further test of this hypothesis, we explored whether changing the stimulus format reduced the CSE more in Experiment 1 than in Experiment 2. Specifically, we conducted a mixed ANOVA on mean RT wherein Experiment (1, 2) served as the across-participants factor and current trial stimulus format (color patch, color word), format transition (repeat, switch), previous trial congruency (congruent, incongruent), and current trial congruency (congruent, incongruent) served as the within-participants factors[14]. To facilitate this across-experiment comparison, we employed the same trial types as in Experiment 1. Critically, we observed a significant four-way interaction, $F(1,116) = 4.69$, $p = 0.032$, $\eta_p^2 = 0.039$ (see Fig. 5). Consistent with the task set hypothesis, changing the stimulus format from one trial to the next reduced the CSE more in Experiment 1 (32 ms) than in Experiment 2 (-1 ms).[15]

---

[14] To ensure we were comparing the same trial types in the across-experiment comparison, we analyzed only the trials wherein the distractor and target appeared in the same stimulus format (i.e., both words or both patches). Also, as in Experiment 1, we analyzed current trial stimulus format only to account for the variance this factor produced.

[15] As stimulus formats independently varied for both the prime and probe in Experiment 1, but not in Experiment 2, overall task difficulty may have been significantly higher in the latter experiment. Contrary to this possibility, however, mean RT was longer (706 ms vs. 659 ms; $t(116) = 2.97$, $p = 0.004$, *Cohen's d* = 0.56) and the congruency

Discussion

The results of Experiment 2 support the task set hypothesis over the attentional reset hypothesis. Specifically, unlike in Experiment 1, alternating the target's stimulus format across consecutive trials did not reduce the CSE. Such findings suggest that participants perceived color words and color patches as belonging to the same task set in Experiment 2 rather than to separate task sets as in Experiment 1. In line with these findings, our exploratory across-experiment comparison further revealed that changing the target's stimulus format reduced the CSE less in Experiment 2 than in Experiment 1. Thus, these findings suggest that task sets based on stimulus formats, rather than stimulus format on its own, serve as boundaries for the CSE in standard visual tasks.

Critically, our findings further suggest that the task-relevance of stimulus format led to color words and color patches being represented separately in Experiment 1 but not in Experiment 2. For instance, in Experiment 1 the distractor's format always predicted the format of the upcoming target. Thus, participants could employ the distractor's format as a useful, task-relevant feature to prepare for the upcoming target. In contrast, in Experiment 2 the distractor's stimulus format did not validly predict the format of the upcoming target. Thus, participants could never employ the distractor's stimulus format to prepare for the upcoming target. As such, participants were more likely to employ stimulus format to form separate, format-specific task sets in Experiment 1 than in Experiment 2.

---

effect was larger (135 ms vs. 90 ms; $t(116) = 5.66$, $p < 0.001$, *Cohen's* $d = 1.06$) in Experiment 1 than in Experiment 2. Mean ER did not significantly differ between the two experiments ($p > 0.20$).

General Discussion

We sought to distinguish between the task set and attentional reset accounts of contextual boundaries for the CSE in a purely visual task. To this end, we employed a prime-probe task wherein participants could (in Experiment 1) or could not (in Experiment 2) use distinct stimulus formats (i.e., color patches and color words) to create format-specific task sets. Consistent with the task set hypothesis, but not with the attentional reset hypothesis, we found that changing the stimulus format reduced the CSE only when participants could create format-specific task sets (i.e., only in Experiment 1). As we describe next, these findings have important implications for our understanding of the contextual boundaries that limit the scope of adaptive control.

Implications for the task set hypothesis

The present findings extend prior data indicating that task sets serve as boundaries for the confound-minimized CSE in a cross-modal task (Grant et al., 2020) to a unimodal visual task. This is important because researchers employ unimodal tasks to investigate the boundaries of adaptive control far more often than they employ cross-modal tasks. Thus, the present findings reveal the nature of CSE boundaries under conditions that are more typical in the literature.

Our findings also speak to the idea that distinct visual features on their own are not sufficiently "salient" to trigger task set formation (Hazeltine et al., 2011, Experiment 2). Hazeltine et al. argued that the task instructions must map the stimuli associated with distinct features to fingers on different hands in order for those features to serve as task set boundaries for the CSE. Doing so makes the features informative regarding the hand participants should use to respond and, hence, sufficiently "salient" (i.e., distinct) to trigger task set formation. Contrary to this view, we found that participants *are* able to form distinct task sets related to distinct visual features if a task maps the stimuli associated with those features to the same fingers. This discrepancy may indicate that the visual features that allow participants to distinguish between different subsets of stimuli are more salient in the present task (i.e., color patches and color words) than in Hazeltine et al.'s task (i.e., central letters and peripheral circles). It may also reflect our use of a confound-minimized protocol or greater statistical power in the present study due to the larger sample size that we employed (i.e., 48 participants vs. 16 participants).

In line with the possibility that some visual features are more salient than others, the results of the exploratory analyses in Experiment 1 suggest that color words are more strongly bound in a memory of the previous trial than color patches, resulting in a larger task set boundary for the CSE. This outcome fits with an emerging view wherein episodic binding and episodic retrieval make separable contributions to adaptive control (Frings et al., 2020). It also suggests that visual features may vary on a continuum of salience, such that task set boundaries for the CSE are easier to create (or observe) for some features than for others. Future studies could investigate what determines whether the visual features that distinguish between different subsets of trials are salient enough to trigger task set formation.

Future studies could also investigate whether orienting attention to the distractor's stimulus format triggers task set formation in the same way that orienting attention to the distractor's sensory modality triggers task set formation in cross-modal tasks (Grant et al., 2020). To investigate this possibility, researchers could employ a study design that is analogous to Experiment 3 in Grant et al.'s (2020) cross-modal study with the exception that the stimuli are color patches and color words in the visual modality, rather than directional words in the visual and auditory modalities. The distractor's format would predict the upcoming target's format about 90% of the time (format valid trials) but not the rest of the time (format invalid trials).

If participants orient attention to the distractor's format to prepare for an upcoming target in the same format, then they should respond more slowly in format invalid trials than in format valid trials. Moreover, this *format validity effect* should be larger than in a design like that of the present Experiment 2, wherein participants are unlikely to orient attention to the distractor's format because it does not consistently predict the target's format. Finally, one should observe format-specific CSEs, although these might be small given the presence of mixed-format trials that discourage orienting to some degree. Consistent with these predictions, we reported exactly analogous effects in our cross-modal tasks (Grant et al., 2020). Most important, observing such effects in a purely visual task would suggest that orienting attention to the stimulus properties of a distractor is a relatively general mechanism for triggering task set formation.

Implications for the attentional reset hypothesis

The present data do not support the attentional reset hypothesis, but it remains possible that boundaries for the CSE in other tasks reflect an attentional reset. This may explain, for example, why changing the gender (male or female) of a spoken auditory distractor word reduces the CSE even though orienting attention to voice gender cannot facilitate the identification of a gender-neutral target (a high- or low-pitched tone) (Spapé & Hommel, 2008). This may also explain why changing irrelevant stimulus colors reduces the CSE even when identifying a target requires attention to a different feature (i.e., stimulus form) (Braem, Hickey, et al., 2014). Indeed, varying the irrelevant stimulus dimension (e.g., stimulus format) often interferes with responding to the relevant stimulus dimension (e.g., color) of a target stimulus (i.e., the Garner interference effect; Garner, 1974). Alternatively, these findings may indicate that participants create format-specific task sets even when orienting attention to the perceptual features of the distractor cannot facilitate the identification of the target. To resolve this ambiguity, future researchers could attempt to distinguish between the attentional reset and task set hypotheses in tasks like those above.

Implications for views wherein adaptive control is dimension-specific

The results of Experiment 2 contradict a central assumption of the conflict-monitoring model (Botvinick et al., 2001; Egner, 2014), which is that adaptive control processes are specific to the task-relevant stimulus dimension. Specifically, we observed a robust CSE even when the task-relevant stimulus dimension changed (i.e., from color patches to color words, or vice-versa). Since the conflict-monitoring model is highly influential in the literature on adaptive control, this outcome may appear surprising. Recent data from other confound-minimized tasks, however, further suggest the possibility that the conflict-monitoring model cannot fully explain the CSE.

For example, the size of the CSE in the prime-probe task does not vary with whether there is a large congruency effect in mean RT or no congruency effect at all (Weissman, Egner, et al., 2015) even though, in the latter case, conflict is unlikely to differ between congruent and incongruent trials (Yeung, Cohen, & Botvinick, 2011). Further, when there is no overall congruency effect, the CSE is associated with a positive congruency effect after congruent trials and with a negative, or *reverse*, congruency effect after incongruent trials (i.e., faster mean RT in incongruent trials than in congruent trials). A reverse congruency effect is consistent with models wherein control processes modulate (e.g., inhibit) the response cued by the distractor after incongruent trials (Ridderinkhof, 2002). However, it does not fit with the conflict-monitoring model wherein control processes shift attention toward the target and away from the distractor after incongruent relative to congruent trials (Botvinick et al., 2001). Indeed, even shifting all of one's attention away from the distractor could eliminate the congruency effect but not reverse it. The present findings, therefore, add to a growing body of data indicating that the conflict-monitoring model does not provide an adequate explanation of the CSE in the prime-probe task.

The results of Experiment 2 also inform inhibitory views of the CSE wherein adaptive control processes are specific to the task-*irrelevant* (i.e., distractor) stimulus dimension (Kim, Lee, & Cho, 2015; Lee & Cho, 2013). Specifically, the size of the CSE did not vary with whether the distractor's stimulus format repeated or switched in consecutive trials. Given prior findings indicating that a modulation (e.g., inhibition) of the response cued by the distractor contributes to the CSE (Ridderinkhof, 2002; Stürmer et al., 2002; Weissman, Egner, et al., 2015), we do not reject the claim that the CSE indexes a modulation of response activation. However, as we explain next, this modulation may occur more strongly when the task set that underlies performance specifies the format in which the distractor appears (Grant et al., 2020).

Specifically, our view is as follows. When participants create format-specific task sets (as in the present Experiment 1), they form separate task sets for color words and color patches. Each of these task sets includes information only about stimuli that appear in a single stimulus format (e.g., color word *or* color patch). Repeating the previous trial's stimulus format in the current trial triggers the retrieval of the previous trial's event file, which includes the control parameters that were employed. Thus, control processes modulate the response cued by the current-trial distractor, thereby engendering a CSE, only when the previous trial's format repeats in the current trial. In contrast, when participants create a format-general task set (as in the present Experiment 2), they form a single task set for color words and color patches. This task set includes information about stimuli that appear in either format (i.e., color word *and* color patch). The appearance of a stimulus in either format in the current trial triggers the retrieval of the previous trial's event file. Thus, control processes modulate the response cued by the distractor regardless of whether its stimulus format matches or mismatches that of the distractor in the previous trial. To test this view, future studies could use force-sensitive keys to determine whether task sets serve as boundaries for finger-specific modulations of response activation that occur after the distractor appears (Weissman, 2019).

Along these lines, one may wonder how our findings relate to those from prior studies investigating the influence of task dimensions on the boundaries of the CSE (Kim, Lee, & Cho, 2015; Lim & Cho, 2020). For instance, consider Experiment 1 of Kim et al. (2015), wherein the task (Stroop or Simon) alternated on every trial. Here, the task-relevant dimension was different in each task (word – Stroop; color – Simon) but the task-irrelevant dimension was always the same (arrow). Notably, the authors observed a significant CSE, suggesting that the task-irrelevant (vs. task-relevant) dimension contributes to the boundaries of the CSE.

At first glance, these prior findings appear to contradict those from the present Experiment 1. Specifically, they suggest that a change to the task-relevant dimension (e.g., word vs. color) does not reduce the CSE, even though a change to this dimension was associated with a reduction of the CSE in the present Experiment 1. However, these findings are consistent with one another if viewed from the perspective of the task-irrelevant dimension. In Experiment 1 of Kim et al. (2015), the task-irrelevant dimension was an arrow in both word (Stroop) and color (Simon) trials. Thus, participants could not use the task-irrelevant dimension to predict whether the task-relevant dimension was a "word" or a "color" and were therefore unlikely to create format-specific task sets. In Experiment 1 of the present study, however, the format of the task-irrelevant dimension predicted the format of the task-relevant dimension. Thus, participants were likely to use the task-irrelevant dimension (i.e., the distractor's stimulus format) to categorize each trial as “color word” or “color patch” and create format-specific task sets. In short, our findings are consistent with prior data suggesting that the control processes underlying the CSE are specific to task-irrelevant stimulus dimensions (Kim et al., 2015; Lim & Cho, 2020). However, they further suggest that task sets based on irrelevant stimulus dimensions, rather than such dimensions on their own, serve as boundaries for the CSE.

Implications for our understanding of “domain-general” adaptive control

The present finding that task sets – rather than stimulus dimensions – determine the scope of adaptive control speaks to a paradox in the literature on adaptive control. As we described earlier, most studies report an elimination of the CSE when salient perceptual features, or even entire tasks, change across trials (Braem, Hickey, et al., 2014; Dignath et al., 2019; Kiesel et al., 2006; Spapé & Hommel, 2008). However, a growing minority report robust CSEs effects even when stimuli and tasks change radically (Adler, Kroff, & Novick, 2020; Hsu & Novick, 2016; Kan et al., 2013). For instance, experiencing conflict during a sentence-reading or perceptual task reduces interference during a subsequent Stroop task (e.g., Kan et al. 2013).

The literature offers three explanations for such “domain-general” CSEs, the first two of which are consistent with the episodic retrieval view. First, domain-general CSEs may occur when task set boundaries are not very salient (e.g., when participants conceptualize sentence-reading trials and color-naming Stroop trials as part of the same task) (Hazeltine et al., 2011). In this scenario, participants form a single task set that includes all of the trials. Therefore, the repetition of any previous-trial feature in the current trial triggers the retrieval of the previous trial’s event file and engenders a CSE. Second, domain-general CSEs may occur when the local temporal context triggers the retrieval of the previous trial’s event file even in the absence of direct stimulus or response repetitions (Egner, 2014; Frings et al., 2020). The idea here is that repeating the previous trial's task set is sufficient on its own to serve as a retrieval cue (Egner, 2014; Frings et al., 2020). Third, domain-general CSEs may occur when two task sets are so different that they can be simultaneously maintained in working memory because they do not interfere with each other (Braem, Abrahamse, et al., 2014). According to this view, maintaining both task sets in working memory, which is possible only when they do not interfere with each other, allows adaptive control processes to transfer from one task set to another.

In our view, the present findings motivate an account of domain-general CSEs wherein participants associate different perceptual features with the same task set. However, they do not rule out the alternative working memory hypothesis described above. Future studies could investigate this hypothesis by determining whether participants simultaneously maintain non-interfering task sets in working memory. If they do, then taxing the working memory processes that maintain either task set in working memory should reduce or eliminate domain-general CSEs (c.f., Soutschek, Strobach, & Schubert, 2013).

Limitations

Although we argued that participants switched between format-specific task sets in Experiment 1, we did not observe an effect of format transition (i.e., switch costs) in the current study. Two factors may account for the absence of switch costs. First, to avoid feature integration confounds, our task design did not permit stimuli and/or responses to repeat across consecutive trials. Therefore, no stimuli or responses from the previous trial (e.g., D key – red, F key – blue) repeated in the current trial (e.g., H key – green, J key – yellow). In the case of a task switch, the absence of such repetitions may reduce the extent to which participants retrieve an irrelevant task representation from the previous trial (Kiesel et al., 2010; Frings et al., 2020). Consistent with this view, preventing such repetitions from occurring reduces task-switch costs (Schmidt & Liefooghe, 2016). This may partly explain why relatively weak (or absent) task-switch costs have been observed in recent studies of the confound-minimized CSE (Grant et al., 2020; Dignath et al., 2019). Second, in the current study participants employed the same S-R mapping for both color-patch trials and color-word trials. Thus, the two task sets may have been less distinct than in typical task-switching paradigms (Cookson et al., 2016; 2019), thereby further reducing task switch costs.

In the absence of such costs, the overall analyses do not provide independent evidence that participants formed distinct, format-specific task sets in Experiment 1. Future studies could investigate whether such evidence appears in other tasks. Regardless of the outcome, however, the present findings clearly show that task structure influences whether or not salient perceptual features serve as boundaries for the CSE, which is highly consistent with the task set hypothesis.

Conclusion

The present findings indicate that task sets based on salient perceptual features, rather than such features on their own, serve as boundaries for the CSE in a visual distractor-interference task. They also suggest that task set boundaries for the CSE are easier to observe for some visual features than for others. These findings extend an emerging episodic retrieval view of the CSE while posing a challenge to views wherein adaptive control processes are specific to the task-relevant stimulus dimension. They also further motivate a task set account of "domain-general" CSEs. Thus, future work investigating the potentially separable contributions of stimulus salience, episodic binding (in the previous trial), and episodic retrieval (in the current trial) to task set formation may provide additional insights into the boundaries of adaptive control.

Chapter 6 The Binary Structure of Event Files Generalizes to Abstract Features: A Non-Hierarchical Explanation of Task Set Boundaries for the Congruency Sequence Effect

Abstract

Current views posit that forming and retrieving memories of ongoing events influences action control. However, the organizational structure of these memories, or event files, remains unclear. The *hierarchical coding view* posits a hierarchical structure, wherein task sets occupy a high level of the hierarchy. Here, the contents of an event file can be retrieved only if the task set repeats. In contrast, the *binary coding view* posits a non-hierarchical structure, which consists of a collection of independent, binary bindings between different feature pairs. In this view, repeating an abstract feature from a previous event (e.g., the previous trial's S-R mapping) triggers the retrieval of the associated feature from the same binding (e.g., the previous trial's congruency) even if the task set changes. To distinguish between these views, we investigated the nature of task set boundaries for the congruency sequence effect (CSE), an index of adaptive control that reflects event file formation and retrieval. Specifically, we investigated whether or not a CSE appears when the task set changes but the previous trial's S-R mapping repeats. Two experiments involving a cross-modal prime-probe task yielded a CSE under these conditions and ruled out alternative explanations. These findings show that the typical binary structure of event files generalizes from concrete features (e.g., colors and locations) to abstract features (e.g., S-R mappings and task sets). Therefore, contrary to the hierarchical coding view, they provide a non-hierarchical explanation of task set boundaries for the CSE.

Introduction

Engaging in purposeful behavior requires adapting to an ever-changing environment. Considered one of the three main cognitive-core functions (Diamond, 2013), such cognitive flexibility requires humans to combine internal information (e.g., task goals) with external contextual information to achieve a goal (e.g., Miller & Cohen, 2001). For instance, consider two pedestrians who are attempting to cross a busy intersection. If the walk sign does not function correctly, the pedestrians must adapt by shifting their attention from the typical source of information (i.e., the walk sign) to a new source (i.e., the traffic light) to decide when it is safe to cross the street. That is, to navigate the intersection safely, the pedestrians must combine their internal goal of crossing the street with external contextual information in a flexible manner.

To investigate how adaptive control processes operate in the laboratory, researchers measure performance (e.g., reaction time and/or error rate) in distractor-interference tasks (Eriksen & Eriksen, 1974; Eriksen & Schultz, 1979; Simon & Rudell, 1967; Stroop, 1935). For instance, consider the prime-probe task, wherein study participants ignore an initial prime (or distractor) and respond to a subsequent probe (or target) in every trial. In general, participants respond more slowly and less accurately when the prime signals a different (incongruent) response than the probe as compared to the same (congruent) response. However, this congruency effect – an index of overall distractibility – is smaller after incongruent trials than after congruent trials (Gratton, Coles, & Donchin, 1992). This *congruency sequence effect*, or CSE, may reflect adaptive control processes that minimize distraction from irrelevant stimuli (Gratton et al., 1992; Botvinick, Braver, Barch, Carter, & Cohen, 2001). In line with this view, a CSE appears in 4-alternative-forced-choice (4-AFC) tasks even without feature integration (i.e., stimulus and response repetition) and contingency learning (i.e., stimulus frequency) confounds that can independently produce this sequential trial effect (Schmidt & Weissman, 2014; Jimenez & Mendez, 2014; Kim & Cho, 2014; Weissman, Jiang, & Egner, 2014).

The attentional shift and response modulation accounts

There are two primary accounts of the CSE in confound-minimized 4-AFC tasks. According to the *attentional shift account*, control processes shift attention toward the target and away from the distractor to a greater extent after incongruent trials than after congruent trials (Botvinick, 2007; Botvinick et al., 2001; Dreisbach & Fischer, 2012; Gratton et al., 1992). According to the *response modulation account*, control processes modulate (i.e., inhibit and/or enhance) the response signaled by the distractor differently after incongruent trials than after congruent trials (Logan, 1985; Logan & Zbrodoff, 1979; Ridderinkhof, 2002; Stürmer, Leuthold, Soetens, Schroter, & Sommer, 2002; Weissman, Colter, Grant, & Bissett, 2017).

Prior findings from the prime-probe task favor the response modulation account over the attentional shift account (Weissman, Egner, Hawks, & Link, 2015). First, the CSE is larger when the prime appears before (versus with) the probe, consistent with the view that control processes engender the CSE by modulating (e.g., inhibiting) the response cued by the prime before participants respond to the probe (c.f., Ridderinkhof, 2002). Second, when a long (i.e., 1000 ms) inter-stimulus-interval (ISI) separates the prime's offset from the probe's onset, which eliminates the overall congruency effect, the CSE is associated with a positive congruency effect after congruent trials and a *negative* congruency effect (i.e., faster response times in incongruent trials than in congruent trials) after incongruent trials. The latter finding is inconsistent with the attentional shift account. Even completely ignoring the prime after incongruent trials could eliminate the congruency effect but not reverse it. In contrast, this finding is consistent with the response modulation account. For example, if participants inhibit the response cued by the prime, response times to the subsequent probe should be slower in congruent trials, wherein the probe requires the same (inhibited) response, than in incongruent trials, wherein the probe requires a different (uninhibited) response, leading to a negative congruency effect. Finally, we note that since a CSE can appear without an overall congruency effect, it is unlikely that conflict drives the CSE. Consistent with this view, the size of the CSE is typically independent of the size of the congruency effect (Weissman et al., 2014; Weissman, Egner, et al., 2015).

The episodic retrieval view of the CSE

More broadly, the findings above are consistent with an emerging episodic retrieval view of the CSE (Dignath, Johannsen, Hommel, & Kiesel, 2019; Egner, 2014; Hazeltine, Lightman, Schwarb, & Schumacher, 2011; Schumacher & Hazeltine, 2016; Spapé & Hommel, 2008). According to this view, participants form an episodic memory of the previous trial that stores both (1) concrete features (e.g., stimuli and responses) and (2) abstract features (e.g., task sets, S-R mappings, control settings, and trial congruency). Repeating a previous-trial feature (e.g., the task set) in the current trial triggers the retrieval of a memory of the previous trial, which specifies its congruency. This biases control processes to treat the current trial as a congruency repetition. For instance, after an incongruent (vs. congruent) trial, control processes become biased to inhibit (vs. enhance) the response associated with the distractor. This leads to a smaller congruency effect after incongruent relative to congruent trials (i.e., a CSE).

Recent data from a *modified* prime-probe task provide novel support for the episodic retrieval view (Grant & Weissman, 2019). As in the *standard* prime-probe task described earlier, a 1000 ms temporal interval separates the prime and probe. Unlike in the standard prime-probe task, however, participants respond to the prime and the probe using the same S-R mapping (e.g., "identify the directional word – Left, Right, Up, or Down – by making one of four possible keypresses"). The CSE in mean probe response time is much larger in the modified prime-probe task than in the standard prime-probe task, wherein participants do not respond to the prime. Further, this increase in CSE magnitude does not occur when participants respond to the prime in only (a) the previous trial or (b) the current trial. These findings support the episodic retrieval view. Indeed, only in the modified prime-probe task does responding to the prime in the current trial constitute a repetition of the previous trial's overall S-R mapping (i.e., a repetition of both the previous prime's S-R mapping and the previous probe's S-R mapping). The repetition of this abstract feature, which is not present in the standard prime-probe task, should serve as an additional retrieval cue for the memory of the previous trial, leading to an especially large CSE.

The hierarchical and binary coding views of the CSE

Although there is a great deal of support for the episodic retrieval account of the CSE, competing views differ with regard to the exact conditions under which repeating an abstract previous-trial feature engenders a CSE in confound-minimized tasks. The *hierarchical coding view* posits that information about the previous trial's congruency, which gives rise to the CSE, is specific to the abstract task set in which it was encountered. Thus, the previous trial's congruency is only retrieved if the task set repeats. This view is based on neurophysiological findings and computational models of prefrontal cortex (e.g., Badre, 2008; Koechlin, 2008; Miller & Cohen, 2001), which suggest that concrete and abstract features are encoded in memory within a hierarchical framework known as a *task file* (Schumacher & Hazeltine, 2016). Here, *lower levels* consist of concrete features such as stimuli and responses while *higher levels* consist of abstract features such as task sets, stimulus-response (S-R) mappings, motivational factors, and task goals (Hazeltine et al., 2011; Schumacher & Hazeltine, 2016; Cookson, Hazeltine, & Schumacher, 2016, 2019).

An example is useful for illustrating how the hierarchical coding view explains task set boundaries for the CSE. Consider a cross-modal prime-probe task wherein the prime and probe words “Left”, “Right”, “Up”, and “Down” appear in the visual modality in 50% of the trials and in the auditory modality in the other 50%. A hierarchical representation of this task would specify high-level visual and auditory task sets that are associated with low-level stimuli in the visual and auditory modalities, respectively (Schumacher & Hazeltine, 2016). Selecting the visual or auditory task set following the onset of a prime in the visual or auditory modality would activate the subset of the S-R mapping that relates (a) each of the four visual target words to a specific response or (b) each of the four auditory target words to a specific response. Critically, information about each trial’s congruency (i.e., whether the distractor and target activate the same or different responses) would be stored within the relevant subset of the S-R mapping, which, in turn, would be nested within the modality-specific task set (Hazeltine et al., 2011). Thus, the CSE would vanish if the task set changes in the next trial, because this would prevent the retrieval of the previous trial’s congruency. In line with this view, switching between modality-specific task sets typically eliminates the CSE (Grant, Cookson, & Weissman, 2020; Hazeltine et al., 2011; Kreutzfeldt, Stephan, Willmes, & Koch, 2016; Yang et al., 2017).

In contrast, the *binary coding view* posits that repeating the previous trial’s task set is just one of many ways to cue the retrieval of the previous trial’s congruency. According to this view, participants form an episodic memory of each trial known as an “event file” (Hommel, 1998). Unlike a task file, an event file consists of a *non-hierarchical* organization of independent, binary bindings between concrete and/or abstract features. Most important for present purposes, the previous trial's congruency, or associated control parameters, can be stored in more than one binding (e.g., a binding with the previous trial's abstract task set, a binding with the previous trial's abstract S-R mapping, etc.) (Dignath et al., 2019). Repeating an abstract feature from the previous trial, such as the S-R mapping, can therefore independently cue the retrieval of the previous trial's congruency even when another abstract feature, such as the task set, changes[16].

In short, unlike the hierarchical coding view, the binary coding view posits that a CSE may appear when the task set changes. A CSE may appear, for example, when the modality-specific task set changes in the cross-modal version of the *modified* prime-probe task described earlier. The reason is that participants can use the same overall (i.e., not hierarchically organized) S-R mapping (i.e., "identify the directional word by making one of four possible keypresses"), regardless of whether the stimuli appear in the visual or the auditory modality. Therefore, responding to the current-trial prime constitutes a repetition of the previous trial's S-R mapping, which should trigger the retrieval of the previous trial's congruency and thereby engender a CSE.

The present study

In the present study, we sought to distinguish between the hierarchical and binary coding views of task set boundaries for the CSE. In Experiment 1, we employed a cross-modal version of the *modified* prime-probe task, wherein participants respond to both the prime and the probe. Consistent with the binary coding view, but not with the hierarchical coding view, we observed a

---

[16] Concrete stimulus and response features can also be stored in multiple bindings that subsequently exert independent influences on performance (Giesen & Rothermund, 2014; Hommel, 1998; Huffman et al., 2020).

significant CSE even when the task set changed. Thus, in Experiment 2 we further explored the binary coding view and ruled out two alternative explanations for our findings in Experiment 1.

Experiment 1

In Experiment 1, we sought to distinguish between the hierarchical and binary coding views by using a cross-modal version of the *modified* prime-probe task. Here, the prime and probe appear in the visual modality in half the trials and in the auditory modality in the other half. This allows participants to categorize each trial as “visual” or “auditory”, enabling them to form modality-specific task sets (Hazeltine et al., 2011). Finally, participants respond to (1) the prime during the 1150 ms inter-stimulus-interval (ISI) that separates prime offset from probe onset and (2) the probe after it appears (during a second ISI). As we described earlier, the hierarchical coding view predicts that changing the modality-specific task set across consecutive trials will eliminate the CSE just as it does in the cross-modal version of the standard prime-probe task (Grant et al., 2020; Hazeltine et al., 2011). In contrast, the binary coding view predicts that a significant CSE may appear even when the modality-specific task set changes.

Methods

Participants

Prior to collecting data, we pre-registered all hypotheses, methods, and analyses on the Open Science Framework, or OSF (https://osf.io/x75nq). As discussed in the pre-registration, we conducted a power analysis (via G*Power 3.1.9.2; Faul, Erdfelder, Lang, & Buchner, 2007) to determine the required sample size for observing an interaction among modality transition, previous trial congruency (i.e., trial *N – 1* congruency), and current trial congruency (i.e., trial *N* congruency). To this end, we employed the effect size for this interaction (0.37) from Hazeltine et al. (2011) and an alpha (α) of 0.05. The results of this analysis indicated that a sample size of 32 participants would allow us to observe this interaction with sufficient power (i.e., greater than 99% power) in the present experiment.

We collected data from thirty-nine students at the University of Michigan, all of whom participated through the Intro Psych Subject Pool for course credit. We excluded all participants who self-reported a visual and/or auditory impairment (1 participant) or completed the task with less than 75% overall accuracy (6 participants). We conducted statistical analyses on the data from the remaining 32 participants (7 male, 25 female; 31 right-handed, 1 left-handed; 18-20 years old, $\bar{x}$ = 18.66, SD = 0.70). No participants reported neurological (e.g., seizures, ADHD) or perceptual (e.g., uncorrected vision) impairments. The University of Michigan's Behavioral Sciences Internal Review Board approved all experimental procedures. The raw data is available via the OSF (https://osf.io/a85bf/?view_only=c5bdeaf264e94370be943e111fb83960).

Task Stimuli and Experimental Design

The primes and probes were four direction words (i.e., left, right, up, and down) that could appear in the visual or auditory modality. The visual angles (visual words) and decibels (auditory words) were the same as those in one of our prior studies (Grant et al., 2020). To present the task stimuli and record participants' responses (via a standard QWERTY keyboard), we employed the Psychophysics Toolbox in MATLAB (Brainard, 1997). Participants viewed all visual stimuli from a distance of approximately 55 cm.

A 2 s fixation cross appeared at the beginning and end of each block. Each trial consisted of four events: (1) a distractor; (2) an inter-stimulus-interval (ISI); (3) a target, and (4) an inter-trial-interval (ITI). The duration of each event is shown in Figure 1. We employed a longer ISI between the prime and the probe than in prior studies (e.g., Hazeltine et al., 2011) to provide participants with enough time to respond to the prime before the probe appeared.

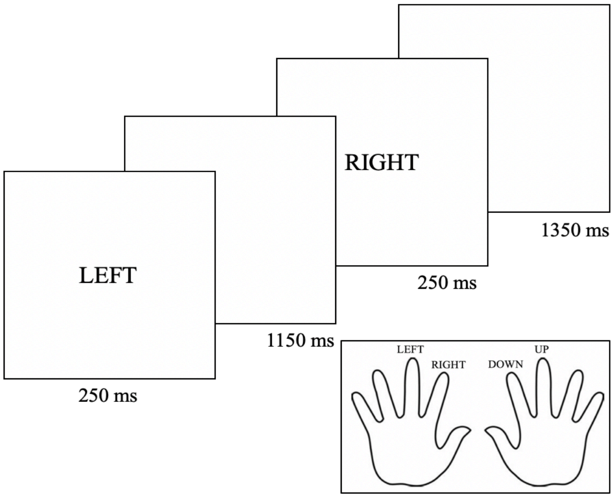


Figure 6-1. The prime-probe task used in Experiment 1 In each trial, a distractor word preceded a target word (the figure illustrates one trial). Trials contained either visual stimuli (shown above) or auditory stimuli. Visual stimuli appeared in white on a black background. Auditory stimuli were delivered via headphones. The number beneath each box indicates the length of the corresponding trial component in milliseconds (ms). The hands indicate the stimulus-response mapping.

There were sixteen prime-probe pairs (eight visual, eight auditory). In each modality (e.g., visual), half of the prime-probe pairs were congruent (e.g., UP-UP) and half were incongruent (e.g., UP-DOWN). Thus, in every block, there were four main trial types: congruent trials in the visual modality, congruent trials in the auditory modality, incongruent trials in the visual modality, and incongruent trials in the auditory modality. Each of these four trial types preceded and followed every trial type roughly equally often, separately for odd and even trials. Specifically, one sequential trial type appeared one less time than the others in each block, because there was no "previous trial type" for the first trial. However, since the trial sequence differed for each block, the underrepresented trial type varied randomly across blocks.

We avoided two prevalent confounds. To avoid feature integration confounds (Hommel, Proctor, & Vu, 2004; Mayr, Awh, & Laurey, 2003), we presented prime-probe pairs involving the words left and/or right in odd trials and the words up and/or down in even trials[17]. To avoid contingency learning confounds (Schmidt & De Houwer, 2011), we paired each prime (e.g., UP) equally often with the congruent (e.g., UP) and incongruent (e.g., DOWN) probe with which it could appear.

Procedure

We instructed participants to identify each word in a given trial by pressing the F key (to indicate left), the G key (to indicate right), the J key (to indicate up), or the N key (to indicate down) (see Fig. 1). If a participant (a) did not respond within 0.900 seconds of prime or probe onset or (b) pressed the wrong key (error trials), an error message appeared on the screen for 200 ms. Participants completed a practice block (64 trials) and twelve subsequent test blocks (64 trials each). Afterwards, a research assistant explained the purpose of the study.

Data Analysis

Before analyzing the data, we excluded several trial types. Prior to analyzing mean probe response time (RT), we excluded practice trials, the first trial of each block, outliers (i.e., trials with RTs > 3 SD from the conditional mean), omitted responses, incorrect responses (i.e., errors), and trials immediately following omitted responses and errors. Prior to analyzing mean

---

[17] One might wonder whether the impact of forming modality-specific task sets on task performance differs between the left/right and up/down stimulus sets. Contrary to this possibility, exploratory analyses of mean probe RT and mean probe ER revealed that switching between (vs. repeating) modality-specific task sets in the present experiments reduced the CSE to the same degree in odd and even trials.

probe error rate (ER), we excluded the same trial types with the exception of errors, which were the dependent measure. In total, 13.2% of the trials were errors and 0.6% were outliers.

Following the exclusions above, we conducted separate, repeated-measures ANOVAs on mean probe RT and mean probe ER. There were four within-participants factors in each ANOVA: current trial modality (visual, auditory), modality transition (repeat, switch), trial *N – 1* congruency (congruent, incongruent), and trial *N* congruency (congruent, incongruent). We included current trial modality to account for the variance this factor produced, but this factor was not of primary interest. For this reason, we do not report our findings related to this factor. However, we note that this factor did not interact with the CSE or with the critical interaction among modality transition, trial *N – 1* congruency, and trial *N* congruency.

Results

Mean Probe RT

There were three significant main effects of modality transition, $F(1,31) = 28.03$, $p < 0.001$, $\eta_p^2 = 0.48$, trial *N* – 1 congruency, $F(1,31) = 22.91$, p < 0.001, $\eta_p^2 = 0.43$, and trial *N* congruency, $F(1,31) = 43.39$, p < 0.001, $\eta_p^2 = 0.58$. These effects occurred because mean probe RT was slower (1) in modality-repeat trials (480 ms) than in modality-switch trials (471 ms), (2) after congruent trials (480 ms) than after incongruent trials (470 ms), and (3) in current incongruent trials (490 ms) than in current congruent trials (460 ms).

There were also two significant two-way interactions. Specifically, there were interactions between modality transition and trial *N* congruency, $F(1,31) = 21.70$, $p < 0.001$, $\eta_p^2 = 0.41$, and between trial *N – 1* congruency and trial *N* congruency, $F(1,31) = 95.31$, $p <$

0.001, η_p^2 = 0.76. These interactions occurred because the congruency effect was (1) larger in modality-switch trials (38 ms) than in modality-repeat trials (23 ms) and (2) smaller after incongruent trials (13 ms) than after congruent trials (49 ms) (i.e., because there was a CSE).

Finally, there was a significant three-way interaction among modality transition, trial *N – 1* congruency, and trial *N* congruency, *F*(1,31) = 20.58, $p < 0.001$, η_p^2 = 0.40 (Fig. 2). Consistent with both the binary and hierarchical coding views, the CSE was larger in modality-repeat trials (48 ms; *F*(1,31) = 85.88, $p < 0.001$, η_p^2 = 0.74) than in modality-switch trials (24 ms; *F*(1,31) = 39.71, $p < 0.001$, η_p^2 = 0.56) (Fig. 2). Inconsistent with the hierarchical coding view, however, the CSE remained significant in modality-switch trials.

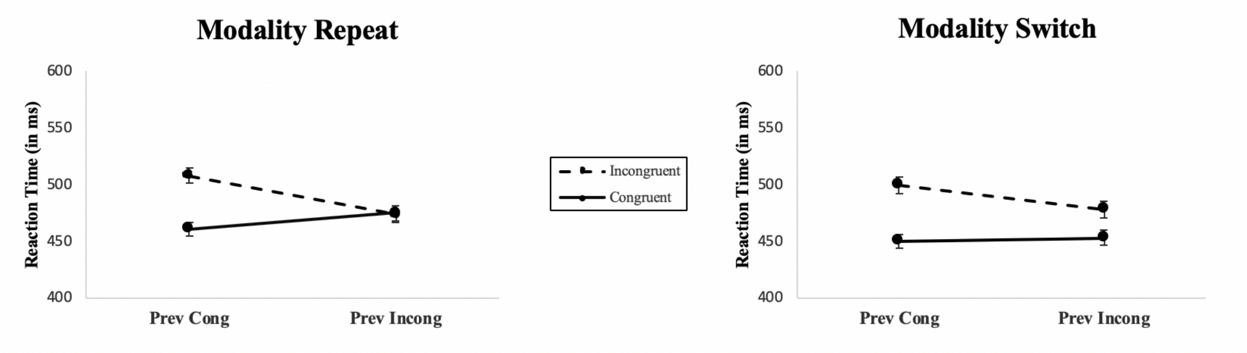


Figure 6-2. The CSE in each of the two main trial types of Experiment 1

Previous trial congruency is indicated on the x-axis (Prev Cong: previous congruent trial; Prev Incong: previous incongruent trial). Current trial congruency is indicated by line type (Incong: dashed line; Cong: black line). Reaction Time (in ms) is indicated on the y-axis. Error bars indicate ±1 standard error of the mean.

Mean Probe ER

There were three significant main effects of modality transition, *F*(1,31) = 12.15, $p = 0.001$, η_p^2 = 0.28, trial *N – 1* congruency, *F*(1,31) = 17.89, $p < 0.001$, η_p^2 = 0.37, and trial *N* congruency, *F*(1,31) = 18.07, $p < 0.001$, η_p^2= 0.37. These main effects occurred because mean ER was higher (1) in modality-repeat trials (11.8%) than in modality-switch trials (10.2%), (2) after congruent trials (12.1%) than after incongruent trials (9.9%), and (3) in current incongruent trials (12.2%) than in current congruent trials (9.8%).

Finally, there was a significant two-way interaction between trial $N - 1$ congruency and trial N congruency, $F(1,31) = 33.48$, $p < 0.001$, $\eta_p^2 = 0.52$. Indicating the presence of a CSE, the congruency effect was smaller after incongruent relative to congruent trials (0.1% vs. 4.6%).

Discussion

We observed a CSE in the modified version of the cross-modal prime-probe task even when the modality-specific task set changed. This result suggests that responding to the prime in the current trial, which constitutes a repetition of the previous trial's S-R mapping, triggers the retrieval of the previous trial's congruency even when participants switch to a different modality-specific task set. This outcome supports the binary coding view.

Mirroring prior findings from the *standard* version of the cross-modal prime-probe task (e.g., Grant et al., 2020; Hazeltine et al., 2011; Kreutzfeldt et al., 2016; Yang et al., 2017), we also observed a smaller CSE in modality-switch trials than in modality-repeat trials. This result suggests that repeating the previous trial's S-R mapping by responding to the prime is not the sole determinant of CSE magnitude. Rather, it would appear that repetitions of the previous trial's (a) modality-specific task set and (b) S-R mapping serve as distinct retrieval cues for a memory of the previous trial, such that removing one of these cues makes episodic retrieval less efficient and reduces the CSE. This interpretation is also consistent with the binary coding view.

However, two alternatives to the binary coding view may account for the significant CSE that we observed when the task set changed. First, the use of a relatively long (i.e., 1150 ms) prime-probe interval may provide participants with enough time to switch attention to the current-trial task set before the probe appears. In particular, it may allow participants to orient attention to the modality of the prime (e.g., auditory), which always predicts the modality of upcoming probe (e.g., auditory), thereby reducing task switch costs that reduce the CSE (Egner, 2008; Kiesel, Kunde, & Hoffmann, 2006). Consistent with this possibility, participants orient attention to the modality in which the prime appears when they know the probe is highly likely to appear in the same modality (Grant et al., 2020). Second, the use of a relatively long prime-probe interval may provide participants with enough time to silently pronounce the prime word before the probe word appears. Thus, participants may use a similar verbal code in visual and auditory trials, leading to a reduced modality-specific boundary for the CSE. In line with this possibility, 1150 ms is enough time to subvocalize one-syllable words (Landauer, 1962; Standing & Curtis, 1989).

Experiment 2

In Experiment 2, we sought to distinguish between the binary coding view and the two alternative explanations of our findings from Experiment 1 described above. To this end, we used a cross-modal version of the *standard* prime-probe task. This task was identical to the task in Experiment 1 with the sole exception that participants responded only to the probe. We reasoned that if using a relatively long prime-probe interval engendered a CSE in Experiment 1 by reducing task-switch costs and/or allowing participants to use the same verbal code in visual and auditory trials, then, as in Experiment 1, changing the task set should not eliminate the CSE.

In contrast, the binary coding view predicts that changing the modality-specific task set will eliminate the CSE in a standard prime-probe task for two reasons. First, the current task set does not cue the retrieval of the previous trial's congruency. Second, withholding a response to the prime in the current trial does not constitute a repetition of the previous trial's overall S-R mapping (i.e., a repetition of both the previous prime's S-R mapping and the previous probe's S-R mapping) that can trigger the retrieval of the previous trial's congruency. In fact, the previous trial does not have an *overall* S-R mapping because participants respond to the probe but not to the preceding prime. Therefore, when the modality-specific task set changes, neither the current task set nor the current S-R mapping can trigger the retrieval of the previous trial's congruency. This should result in the absence of a CSE. Consistent with this prediction, the CSE is smaller in the standard version of the unimodal (i.e., visual-modality) prime-probe task, wherein the overall S-R mapping does not repeat, than in the modified version of the same task, wherein the overall S-R mapping does repeat (Grant & Weissman, 2019).

Methods

Participants

All of our hypotheses, methods, and analyses were pre-registered on the OSF (https://osf.io/x9n36). To determine the appropriate sample size, we conducted two power analyses (G*Power 3.1.9.2; Faul et al., 2007). First, we conducted a power analysis to estimate the sample size needed to observe the critical three-way interaction from Experiment 1. Second, we conducted a power analysis to estimate the sample size needed to observe a CSE in modality-switch trials. To conduct these analyses, we used the effect sizes for the corresponding three-way (η_p^2= 0.40) and two-way (η_p^2 = 0.56) interactions in Experiment 1 and an alpha (α) of 0.05. The results indicated that a sample size of 32 participants would allow us to observe each of these interactions with sufficient power (i.e., greater than 99% power).

Thirty-eight students from the University of Michigan participated through the Intro Psych Study Pool for course credit. We excluded participants who did not complete the task (1), did not follow the task instructions (1), experienced technical (i.e., computer-related) difficulties (1), and/or performed the task with less than 75% accuracy (3). We conducted our analyses on data from the remaining 32 participants (16 male, 16 female; 31 right-handed, 1 left-handed; 17-26 years old, $\bar{x}$ = 18.66, SD = 1.49). No participants reported impairments of a neurological (e.g., seizures, ADHD) and/or perceptual (e.g., uncorrected vision) nature. The University of Michigan's Behavioral Sciences Internal Review Board approved all procedures. The raw data is available via the OSF (https://osf.io/gy7rj/?view_only=70b65aa28dfb4d2a9454d685a959e6fd).

Task Stimuli and Experimental Design

The stimuli, apparatus, and experimental design were identical to those in Experiment 1.

Procedure

The procedure was the same as that in Experiment 1 with one exception: rather than responding to both the prime and the probe, participants responded only to the probe.

Data Analysis

The data analysis was identical to that in Experiment 1. On average, 5.9% of the trials were errors and 0.8% were outliers.

Results

Mean Probe RT

There was a significant main effect of modality transition, $F(1,31) = 57.76$, $p < 0.001$, $\eta_p^2 = 0.65$ and a significant two-way interaction between trial $N - 1$ congruency and trial N congruency, $F(1,31) = 31.20$, $p < 0.001$, $\eta_p^2 = 0.50$. These findings reflect (a) longer RT in modality switch trials than in modality repeat trials (548 ms vs 522 ms) and (b) a CSE: a smaller congruency effect after incongruent relative to congruent trials (-4 ms vs 15 ms).

Finally, there was a significant three-way interaction among modality transition, trial $N - 1$ congruency, and trial N congruency, $F(1,31) = 15.63$, $p < 0.001$, $\eta_p^2 = 0.34$. The CSE was significant in modality-repeat trials (33 ms; $F(1,31) = 42.48$, $p < 0.001$, $\eta_p^2 = 0.58$) but not in modality-switch trials (6 ms; $F(1,31) = 1.69$, $p = 0.20$, $\eta_p^2 = 0.05$) (Fig. 3). Thus, unlike in Experiment 1, a change in modality across consecutive trials completely eliminated the CSE.

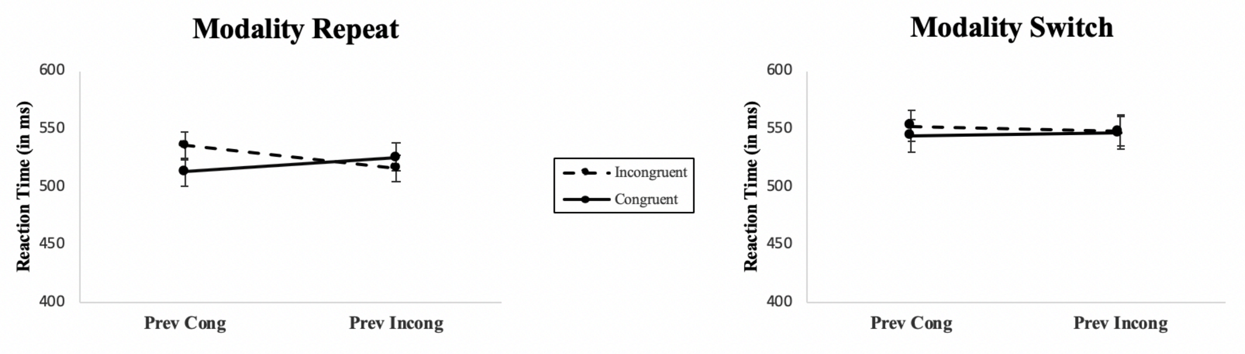


Figure 6-3. The CSE in each of the two main trial types of Experiment 2

Previous trial congruency is indicated on the x-axis (Prev Cong: previous congruent trial; Prev Incong: previous incongruent trial). Current trial congruency is indicated by line type (Incong: dashed line; Cong: black line). Reaction Time (in ms) is indicated on the y-axis. Error bars indicate ±1 standard error of the mean.

Mean Probe ER

There were no significant effects (all p values > 0.25).

Discussion

In line with the binary coding view, we found that switching between task sets eliminates the CSE in a cross-modal version of the *standard* prime-probe task. This task matched the task from Experiment 1 with the exception that participants did not respond to the prime. Thus, our finding suggests that the use of a long (i.e., 1150 ms) prime-probe interval in Experiment 1 cannot explain the CSE that we observed in modality-switch trials. More broadly, our findings in Experiments 1 and 2 suggest that the absence of a CSE in modality-switch trials of the standard prime-probe task (Grant et al., 2020; Hazeltine et al., 2011; Kreutzfeldt et al., 2016; Yang et al., 2017) indexes the cumulative effects of switching between (a) task sets and (b) S-R mappings.

Exploratory Analysis

Our findings thus far suggest that task sets and S-R mappings both serve as boundaries for the CSE. However, two issues remain unclear. First, it remains unclear whether switching between S-R mappings reduces CSE magnitude in the cross-modal prime-probe task as it does in the visual prime-probe task (e.g., Grant & Weissman, 2019). Second, it remains unclear whether task sets and S-R mappings serve as independent or overlapping boundaries for the CSE.

To begin to address these issues, we conducted an exploratory mixed ANOVA wherein experiment (1, 2) served as the between-participants factor, current trial modality (visual, auditory) and modality transition (repeat, switch) served as the within-participants factors, and the CSE in mean probe RT served as the dependent measure. We made two predictions. First, if

S-R mappings serve as boundaries for the CSE, then overall CSE magnitude should be lower in Experiment 2 than in Experiment 1. Second, if task sets and S-R mappings serve as independent (overlapping) boundaries for the CSE, then the influences of experiment and modality transition on CSE magnitude should sum additively (interact).

There were two significant main effects (Fig. 4). First, as in the overall analyses, there was a main effect of modality transition, $F(1,62) = 34.87$, $p < 0.001$, $\eta_p^2 = 0.31$, because the CSE was smaller in modality-switch trials than in modality-repeat trials. Second, there was a main effect of experiment, $F(1,62) = 10.25$, $p = 0.002$, $\eta_p^2 = 0.14$, because the CSE was smaller in Experiment 2 than in Experiment 1. This result conceptually replicates our prior finding that making the prime task-irrelevant (vs. task-relevant) reduces the CSE (Grant & Weissman, 2019).

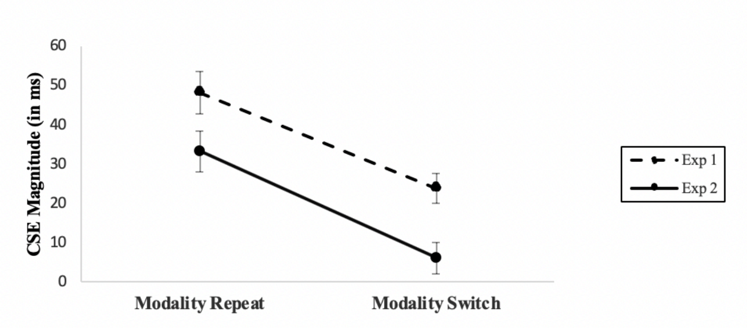


Figure 6-4. CSE magnitude across Experiments 1 and 2

The dashed line indicates the CSE in Experiment 1. The solid line indicates the CSE in Experiment 2. Modality condition (repeat, switch) is indicated on the x axis. CSE magnitude (in ms) is indicated on the y-axis.

Finally, experiment and modality transition did not interact, $F < 1$. Consistent with the binary coding view, this outcome suggests that task sets and S-R mappings serve as independent boundaries for the CSE. Since this finding comes only from an exploratory analysis, however, future studies will be needed to test this hypothesis in an *a priori* fashion.

General Discussion

We sought to distinguish between the hierarchical and binary coding views of task set boundaries for the CSE. Consistent with both views, we found that switching between modality-specific task sets reduces the CSE in both the modified and standard version of the cross-modal prime-probe task. Critically, however, we also found that switching between task sets does not eliminate the CSE in the modified version of the cross-modal prime-probe task. As we explained earlier, this outcome shows that repeating the previous trial's S-R mapping triggers the retrieval of the previous trial's congruency even when the modality-specific task set changes. Therefore, this outcome favors the binary coding view over the hierarchical coding view. It also has important implications for our understanding of task set boundaries for adaptive control.

Implications for the binary coding view

To our knowledge, the present findings are the first to favor the binary coding view over the hierarchical coding view. This outcome extends current views of action control such as the theory of event coding (Hommel, 2001) and the Binding and Retrieval in Action Control (BRAC) framework (Frings et al., 2020). The BRAC framework, for instance, focuses on how the formation and retrieval of binary bindings related to *concrete* features (e.g., stimuli and responses) - rather than *abstract* features (e.g., task sets and S-R mappings) - drives sequential-trial effects such as the CSE, task-switch costs, and negative priming.

For example, consider a standard 2-AFC flanker task involving the letters A and B wherein repetitions of stimuli and/or responses can occur in consecutive trials. Cognitive control accounts of the CSE posit that a previous incongruent trial (e.g., ABA) trial triggers an upregulation of control processes that minimize distraction in the current incongruent trial (e.g., ABA), thereby reducing the congruency effect (e.g., Botvinick et al., 2001; Gratton et al., 1992; Egner, 2008). In contrast, the BRAC framework posits that repeating a previous-trial stimulus feature (e.g., the distractor letter A) triggers the retrieval of the previous trial's response, which is stored in the same binary binding. Critically, retrieving this response facilitates performance to a greater extent when trial congruency repeats than when it alternates, leading to a CSE.

Frings et al. (2020) suggest that the BRAC framework can also account for the CSE when no stimuli or responses repeat across consecutive trials. They posit that repeating the abstract task set or context cues the retrieval of the previous-trial control state. But the manner in which binding and retrieval operate for such abstract features remains unclear. Along these lines, the present findings suggest that the binary structure of event files extends to abstract features such as task sets and S-R mappings. First, we observed a robust CSE in the modality-switch trials of Experiment 1, suggesting that repeating the previous trial's S-R mapping triggers the retrieval of the previous trial's congruency even when the task set changes. Second, the exploratory analyses revealed additive effects of switching task sets and switching S-R mappings on CSE magnitude, further suggesting that task sets and S-R mappings are stored in independent, binary bindings with the previous trial's congruency. These findings extend the BRAC framework by showing that the formation and retrieval of binary bindings between *abstract* features can lead to a CSE.

It is important to note, however, that some findings suggest that bindings between abstract features are more complex than what the binary coding view assumes. For instance, consider findings from a recent study of negative priming, which is a phenomenon wherein the appearance of the previous-trial distractor as the current-trial target leads to performance costs (Mayr, Moller, & Buchner, 2018). In this prior study, the authors found that the previous response was retrieved when (a) the previous-trial distractor appeared as the current-trial target *and* (b) the broader auditory context (i.e., a 300- or 700-Hz sine tone played concurrently with the other stimuli) repeated. Such retrieval did not occur, however, when only the broader auditory context repeated. In other words, repeating the context on its own did *not* lead to the retrieval of the previous-trial distractor response. As Mayr et al. (2018) note, these findings suggest that the structure of an event file may sometimes be more complex than the binary structure that is proposed by current views of action control (Hommel, 2004; Frings et al., 2020). That is, bindings involving abstract features may not always be binary. We discuss this possibility further in the next section.

Implications for the hierarchical coding view

Although one might interpret the present findings as evidence against the view that tasks are hierarchically organized (Schumacher & Hazeltine, 2016), they may simply indicate that tasks are hierarchically organized in some situations but not others. Along these lines, it is useful to consider the influence of task structure on whether a task set is organized hierarchically or non-hierarchically. In particular, the findings from two recent studies suggest that a task set is organized hierarchically when the stimuli that are associated with each of two categories (e.g., faces vs. places) are mapped to responses on different hands rather than on the same hand as in the present study (Cookson, Hazeltine, and Schumacher, 2016; 2019).

In each of these prior studies, participants were assigned to one of two conditions. In Condition 1, four face and four place stimuli were mapped to four fingers on the left and right hands, respectively. The authors argued that this task structure allows participants to form a hierarchical task representation wherein each stimulus category is associated with a different hand. Thus, cueing the higher-level stimulus category (face or place) should allow participants to prepare for the cued category and, subsequently, to prepare to respond with a particular hand. In Condition 2, two face and two place stimuli were mapped to four alternating fingers on each hand. The authors argued that this task structure does not allow participants to form a hierarchical task representation because each stimulus category is associated with responses on each hand. Most important for present purposes, in each study the authors found that cueing the stimulus category leads to hand-specific response preparation in Condition 1 but not Condition 2. Therefore, they argued that task structure plays an important role in determining whether or not a task set is organized hierarchically (Cookson et al., 2016; 2019).

These findings may explain why we observed evidence for the binary coding view in the present study rather than for the hierarchical coding view. Specifically, analogous to Condition 2 in the studies described above, we did not associate the stimuli in each sensory modality with responses on different hands. Rather, participants used the same fingers on each hand to indicate whether the prime and/or probe in every trial was “Left”, “Right”, “Up”, or “Down”, regardless of the sensory modality. Given the findings described above, this task structure may have led participants to employ a non-hierarchical (vs. hierarchical) task representation. To investigate this possibility, future studies could determine whether participants employ a hierarchical task representation in studies of the CSE when different stimulus categories are mapped to different hands (e.g., Hazeltine et al., 2011, Experiment 3). If they do, then, unlike in the present study, a CSE should appear only when the task set repeats in consecutive trials.

Broader implications

The present findings extend prior data suggesting distinct boundaries for the CSE, such as tasks sets (Grant et al., 2020; Hazeltine et al., 2011), conflict types (Egner, 2008; Schlaghecken & Maylor, 2020; but see Weissman, 2020), salient contextual features (Spapé & Hommel, 2008; Braem, Hickey, Duthoo, & Notebaert, 2014), and S-R mappings (Grant & Weissman, 2019) (for a review, see Braem, Abrahamse, Duthoo, & Notebaert, 2014). Specifically, our findings show that the typical elimination of the CSE in modality-switch trials of the standard cross-modal prime-probe task (Grant et al., 2020; Hazeltine et al., 2011; Yang et al., 2017) reflects the cumulative effect of crossing two boundaries - one related to task sets and the other related to S-R mappings - rather than a single boundary related to task sets. Furthermore, our findings suggest that task sets and S-R mappings serve as independent boundaries for the CSE.

The present findings also provide additional support for an emerging view wherein the control processes underlying the CSE contribute to cognition in ways that are broader than minimizing distraction from irrelevant stimuli (Grant & Weissman, 2019; Weissman, 2019). In particular, we observed a CSE in the modified prime-probe task of Experiment 1 even though the prime and probe were both task-relevant. This outcome is consistent with our prior suggestion that adaptive control processes engender a CSE by (1) forming an implicit expectation that the congruency of the previous trial (e.g., congruent) will repeat in the current trial (Egner, 2014) and (2) combining this expectation with the identity of the current-trial prime (e.g., "Left") to prepare for an expected probe (e.g., "Left") (Weissman et al., 2017; Grant & Weissman, 2019). Such a "memory-prediction" algorithm may facilitate task performance when trial congruency repeats (vs. alternates) regardless of whether the prime is task-irrelevant or task-relevant. This may explain why adaptive control processes can engender a CSE even in the absence of task-irrelevant stimuli (Weissman, Grant, & Jones, 2020).

Limitations

One may wonder why we observed better performance in modality-switch trials than in modality-repeat trials in the modified prime-probe task of Experiment 1. Indeed, performance is usually worse (not better) when two trials involve different task sets than when they involve the same task set (Rogers & Monsell, 1995). While this result is atypical, it is important to consider that responding to the prime in modality-switch trials requires participants to switch between modality-specific task sets long before the probe appears. Thus, it may eliminate probe-related switch costs. Switching task sets when the prime appears - rather than repeating the same task set - likely also increases arousal, which may further facilitate performance in modality-switch trials (Rogers & Monsell, 1995). Consistent with this potential explanation, we observed the typical costs of switching between task sets in mean probe RT in the standard prime-probe task of Experiment 2, wherein participants responded only to the probe.

One may also wonder whether it is the heightened task-relevance of the prime, rather than the repetition of the previous trial's S-R mapping, that leads to a larger CSE in the modified version of the cross-modal prime-probe task than in the standard version. The fact that the prime is task-relevant may increase the degree to which its relationship to the probe (i.e., congruent or incongruent) is (a) encoded in the previous trial (c.f., Hommel, 2007) and (b) retrieved in the current trial (c.f., Huffman et al., 2020), thereby increasing the size of the CSE. We cannot rule out this possibility. However, the CSE is not larger in the modified (vs. standard) version of the visual prime-probe task when the prime is task-relevant but associated with a different S-R mapping than the probe (Weissman et al., 2017, Experiment 2). This finding suggests that repeating the previous trial's S-R mapping in the modified version of the cross-modal prime-probe task increases the CSE independent of the prime's task-relevance.

Finally, one may wonder whether boundaries for the CSE that are based on salient stimulus properties appear only in cross-modal tasks. Indeed, unlike a change in learned stimulus categories (Hazeltine et al., 2011) or learned S-R mappings (Lim & Cho, 2018), a change in sensory modality constitutes a qualitative change in the brain regions underlying performance (e.g., auditory vs visual cortex) that may exert a large impact on performance. Consistent with this possibility, even changes to conceptual representations that involve different sensory modalities (e.g., “the light is flickering” vs. “the sound is echoing”) incur task switch costs (Scerrati et al., 2015; Pecher, Zeelenberg, & Barsalou, 2004). As we mentioned earlier, however, boundaries for the CSE that are based on salient stimulus properties also appear in visual tasks (Dignath et al., 2019). Therefore, such boundaries are not specific to cross-modal task protocols.

Conclusion

We sought to distinguish between the hierarchical and binary coding views of task set boundaries for the CSE. Consistent with the binary coding view, but not with the hierarchical coding view, repeating the previous trial’s S-R mapping engendered a CSE even when the task set changed. Further, exploratory analyses revealed additive effects of switching task sets and switching S-R mappings on CSE magnitude. These findings suggest that task sets and S-R mappings independently trigger the retrieval of the previous trial’s congruency. Thus, they extend current views of action control by showing that the binary structure of bindings for concrete features (i.e., stimuli and responses) in event files generalizes to abstract features (e.g., task sets and S-R mappings). Future studies investigating how the structure of episodic memory influences the CSE may provide additional insights into the nature and scope of adaptive control.

Chapter 7 General Conclusions

The studies in this dissertation investigated the boundaries of the CSE. The first two studies suggest that task-relevant features engender (Chapter 2) or amplify (Chapter 3) the CSE. Further, the next two studies suggest that task sets, rather than episodic contexts, serve as boundaries for this sequential effect (Chapters 4 and 5). Lastly, the final study suggests that binary bindings provide a non-hierarchical explanation for task set boundaries of the CSE (Chapter 6). Together, these findings provide a comprehensive overview of the boundaries of adaptive control.

Contingent attentional capture triggers the CSE

Chapter 2 investigated two potential mechanisms for engendering the CSE. According to the contingent attentional capture hypothesis, a distractor should engender a CSE when it shares features in common with a potential target, thereby capturing attention (Moore & Weissman, 2010; Schmidt & Weissman, 2015). According to the perceptual conflict hypothesis, however, a distractor should engender a CSE when it perceptually conflicts with the target (Notebaert & Verguts, 2006). Consistent with the former hypothesis, but inconsistent with the latter, we observed cross-modal CSEs only when the distractor possessed target-defining features.

One may wonder, however, whether an alternative explanation could account for some of these findings. For instance, rather than forming separate attentional sets for word and arrow targets in Experiment 2, one possibility is that varying the target format biased participants *against* forming separate attentional sets. Consistent with this hypothesis, independently varying contextual features across trials eliminates the formation of task-set specific CSEs (Chapter 4). Thus, varying target format across trials may have biased participants to perceive both word and arrow targets as belonging to a single attentional set in Experiment 2.

To investigate this possibility, one could employ an all-visual prime-probe task wherein the distractor format is a word in the majority of trials and the target format (word or arrow) varies across trials. Further, in a minority (e.g., 33%) of trials the format of the distractor is an arrow rather than a word. According to the contingent attentional capture hypothesis, the word distractor should capture attention because it possesses target-defining features in the majority of trials. Thus, the CSE should be larger in "word-distractor" trials than in "arrow-distractor" trials. If, however, varying the target format biases participants against forming separate attentional sets, then the CSE should not significantly differ between these two trial types.

Episodic retrieval amplifies the CSE

Chapter 3 followed up on Chapter 2 by distinguishing between two potential mechanisms that could produce a maximal CSE. In this study, we investigated why the size of the CSE is at its maximum when both the prime and the probe are task-relevant in a small percentage of consecutive "catch" trials (Experiment 1, Weissman et al., 2017). According to the unexpected event hypothesis, these trials constitute an unexpected event which, upon their appearance, triggers an upregulation in proactive inhibitory control processes (Wessel & Aron, 2017). After an incongruent (vs. congruent) "catch" trial, this upregulation in control subsequently enhances any ongoing inhibition of the response signaled by the distractor, thereby enhancing the size of the CSE. According to the episodic retrieval hypothesis, however, consecutive "catch" trials constitute an event wherein salient contextual features (i.e., the S-R mapping) repeat within and across trials. This repetition in episodic context should subsequently amplify the retrieval of previous-trial control processes, leading to an especially large CSE. Consistent with this latter hypothesis, the CSE was largest in consecutive "catch" trials even when these trials were not unexpected events.

Although our findings weigh against the unexpected event hypothesis, inhibitory processes associated with unexpected events may still amplify the CSE. For instance, in studies of inhibitory control, researchers often employ tasks such as the Stop-Signal task (Logan, 1994; Verbruggen & Logan, 2008), wherein participants are instructed to make a response to a target (e.g., "A") in each trial unless an infrequent stop signal is presented (e.g., an auditory tone). Notably, prior findings suggest that processes related to such response inhibition are intrinsically linked to those that minimize distraction (Bissett et al., 2017). This is a critical point, as the unexpected consecutive "catch" trials in Weissman et al. (2017) do not appear to involve control processes related to either inhibition or distraction. Indeed, because both the prime and the probe are targets, there is no need to minimize distracting stimuli or inhibit a response.

To re-visit this question of how unexpected events influence the CSE, one could employ a paradigm similar to that in Experiment 1 of Weissman et al. (2017). As in the original study, "catch" trials and consecutive "catch" trials would occur in 33% and in 10% of trials, respectively. And, as in the original study, a yellow prime arrow would appear in these "catch" trials. In contrast to this study, however, in "catch" trials participants would not respond to the distractor and make the *opposite* response to the target (e.g., responding "right" to a left-pointing arrow). As such, they would need to minimize distraction from the yellow distractor arrow. Further, they would need to employ inhibitory processes to overcome the pre-potent response of indicating the direction in which the target arrow points. Thus, as in more "typical" studies of inhibitory control, catch trials and consecutive catch trials would involve control processes related to both minimizing distraction and response inhibition.

Task sets influence the boundaries of the CSE

In Chapter 4 we sought to determine whether task sets or episodic contexts determine the boundaries of the CSE. According to the task set hypothesis, a change in contextual features (e.g., sensory modality) should reduce the CSE only when these features are tied to different task sets (Hazeltine et al., 2011; Schumacher & Hazeltine, 2016). According to the attentional reset hypothesis, however, a change in contextual features should indicate the beginning of a new episodic event, thereby rendering the previous-trial control processes irrelevant. Thus, a change in contextual features should always reduce the CSE (Kreutzfeldt et al., 2016).

To distinguish between these two hypotheses, in Chapter 4 we employed a cross-modal prime-probe task wherein the sensory modality of the distractor and target varied across trials. Consistent with the task set hypothesis, but inconsistent with the attentional reset hypothesis, a change in sensory modality reduced the CSE only when the overall task structure allowed participants to employ this feature to form modality-specific task sets. These findings suggest that task sets determine the boundaries of the CSE.

We followed up on these findings in Chapter 5, wherein we investigated whether such boundaries of the CSE would occur under unimodal (as opposed to cross-modal) conditions. To this end, we employed the same task design as in the previous Chapter but employed all-visual stimuli as in standard distractor-interference tasks. Consistent with our findings from Chapter 4, a change in stimulus format (i.e., color words or color patches) reduced the CSE only when this feature was tied to separate, format-specific task sets. These findings extend prior support for the task set hypothesis from somewhat atypical cross-modal tasks to standard unimodal tasks. Further, they challenge other views (e.g., the attentional reset hypothesis) that posit a change in salient perceptual features should always result in a boundary for the CSE.

Notably, in these two chapters orienting towards the distractor's contextual features was thought to play a large role in task set formation. However, other findings suggest that orienting is not the only mechanism by which task sets may form. As one primary example, Spapé & Hommel (2008) found that a change in contextual features (i.e., voice gender) eliminated the CSE, despite the distractor not possessing target-defining features. Thus, participants were unable to orient to the distractor's features to prepare for the upcoming target. Critically, however, feature integration processes were confounded with trial congruency in this task. As such, one avenue for future research would be to determine whether feature repetitions or control processes drove the formation of the CSE boundaries that Spapé & Hommel (2008) observed.

Binary coding influences the boundaries of the CSE

The findings from Chapters 4 and 5 support the task set hypothesis. Thus, in Chapter 6 we sought to further investigate the task set hypothesis in the context of episodic retrieval and action control. Currently, two primary views posit that forming and retrieving memories of ongoing events influences action control. However, the organizational structure of these memories, or event files, remains unclear. According to the hierarchical coding hypothesis, various trial features are encoded within a hierarchical structure in episodic memory, with task sets composing the highest level (Schumacher & Hazeltine, 2016). Changing the task set should therefore prevent the retrieval of previous-trial control processes, reducing the CSE. According to the binary coding hypothesis, however, various trial features are encoded in a non-hierarchical manner in episodic memory, such that individual features (e.g., S-R mappings) independently trigger the retrieval of previous-trial control processes (Huffman et al., 2020; Frings et al., 2020). As such, a robust CSE should be observed even when the modality-specific task set changes.

To distinguish between these two possibilities, we employed a cross-modal prime-probe task (as in Chapter 4) but had participants respond to both the prime and probe (Experiment 1) or only to the probe (Experiment 2) in each trial. Consistent with the binary coding hypothesis, we observed a robust CSE when the task set changed but the previous trial's S-R mapping repeated. Thus, contrary to the hierarchical coding view, these findings provide a non-hierarchical explanation of task set boundaries for the CSE.

Broader Implications

At the broadest level, the present findings extend our understanding of how memory and adaptive control processes interact during goal-oriented behavior. For instance, our findings suggest that individuals are more likely to retrieve the memory of previous-trial control parameters under similar (vs. dissimilar) task conditions. Moreover, our findings suggest that whether individuals encode task features in a hierarchical or non-hierarchical manner has a powerful influence on the boundaries of adaptive control.

Considering the above, the present findings also provide a strong basis for investigating memory and attentional processing under a variety of contexts. For instance, researchers could employ the paradigms used in the current study to investigate which aspects of episodic memory drive age-related differences in adaptive control. As one example, researchers could employ the prime-probe paradigm from Chapter 2, which allows one to distinguish between enhanced encoding (i.e., trials wherein the S-R mapping repeats for the prime and probe) and enhanced retrieval (i.e., trials wherein the distractor's S-R mapping repeats in the current trial) within the task. Similarly, future studies could employ the cross-modal prime-probe task from Chapter 4 to investigate whether task-set boundaries of the CSE index healthy aging in adaptive control. In Chapter 4, younger adults formed separate task-sets when the overall task structure allowed them to do so. Thus, the formation of task-set boundaries for the CSE in healthy aging adults would reflect the maintenance of memory and adaptive control processes.

Conclusion

The present findings provide a comprehensive overview on the boundaries of adaptive control. They also extend an emerging episodic retrieval view of the CSE, highlighting the critical influence of memory on adaptive behavior. Further work on the interactions between attention and memory may provide additional insights into the boundaries of adaptive control.

Bibliography

Adler, R. M., Valdés Kroff, J. R., & Novick, J. M. (2020). Does integrating a code-switch during comprehension engage cognitive control? *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *46*(4), 741.

Alexander, W. H., & Brown, J. W. (2011). Medial prefrontal cortex as an action-outcome predictor. *Nature Neuroscience*, *14*, 1338–1344.

Alexander, W. H., & Brown, J. W. (2014). A general role for medial prefrontal cortex in event prediction. *Frontiers in Computational Neuroscience*, *8*, 69.

Altmann, E. M. (2011). Testing probability matching and episodic retrieval accounts of response repetition effects in task switching. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *37*(4), 935.

Anderson, B. A. (2015a). The attention habit: How reward learning shapes attentional selection. *Annals of the New York Academy of Sciences*, *1369*(1), 24-39.

Anderson, B. A. (2015b). Social reward shapes attentional biases. *Cognitive Neuroscience*, *7*(1-4), 30-36.

Anderson, B. A., Folk, C. L., Garrison, R., & Rogers, L. (2016). Mechanisms of habitual approach: Failure to suppress irrelevant responses evoked by previously reward-associated stimuli. *Journal of Experimental Psychology: General*, *145*(6), 796.

Ariga, A., & Yokosawa, K. (2008). Contingent attentional capture occurs by activated target congruence. *Perception & Psychophysics*, *70*(4), 680-687.

Badre, D. (2008). Cognitive control, hierarchy, and the rostro–caudal organization of the frontal lobes. *Trends in Cognitive Sciences*, *12*(5), 193-200.

Bisset, P.G., Grant, L.D., Weissman, D.H. (2017). Resisting distraction and response inhibition trigger similar enhancements of future performance. *Acta Psychologica*, *180*, 40-51.

Botvinick, M. M. (2007). Conflict monitoring and decision making: Reconciling two perspectives on anterior cingulate function. *Cognitive, Affective & Behavioral Neuroscience*, *7*, 356–366.

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.

Botvinick, M., Nystrom, L. E., Fissell, K., Carter, C. S., & Cohen, J. D. (1999). Conflict monitoring versus selection-for-action in anterior cingulate cortex. *Nature, 402*, 179-181.

Braem, S., Abrahamse, E. L., Duthoo, W., & Notebaert, W. (2014). What determines the specificity of conflict adaptation? A review, critical analysis, and proposed synthesis. *Frontiers in Psychology*, *5*, 1134.

Braem, S., Hickey, C., Duthoo, W., & Notebaert, W. (2014). Reward determines the context-sensitivity of cognitive control. *Journal of Experimental Psychology: Human Perception and Performance*, *40*(5), 1769.

Braem, S., Verguts, T., & Notebaert, W. (2011). Conflict adaptation by means of associative learning. *Journal of Experimental Psychology: Human Perception and Performance, 37*, 1662-1666.

Brainard, D. H. (1997). The psychophysics toolbox. *Spatial Vision*, *10*, 433–436.

Braun, C. C., & Silver, N. C. (1995). Interaction of signal word and colour on warning labels: differences in perceived hazard and behavioural compliance, *Ergonomics*, *38*(11), 2207-2220.

Cookson, S. L., Hazeltine, E., & Schumacher, E. H. (2016). Neural representation of stimulus-response associations during task preparation. *Brain Research*, *1648*, 496-505.

Cookson, S. L., Hazeltine, E., & Schumacher, E. H. (2019). Task structure boundaries affect response preparation. *Psychological Research*, 1-12.

Diamond, A. (2013). Executive functions. *Annual Review of Psychology*, *64*, 135–168.

Dignath, D., Johannsen, L., Hommel, B., & Kiesel, A. (2019). Reconciling cognitive-control and episodic-retrieval accounts of sequential conflict modulation: Binding of control-states into event-files. *Journal of Experimental Psychology: Human Perception and Performance*, *45*(9), 1265.

Dreisbach, G., & Fischer, R. (2012). Conflicts as aversive signals. *Brain and Cognition*, *78*(2), 94-98.

Dunlap, G. L., Granda, R. E. & Kustas, M. S. (1986). Observer perceptions of implied hazard: safety signal word and color words. IBM Technical Report T00.3428, IBM.

Duthoo, W., Wühr, P., & Notebaert, W. (2013). The hot-hand fallacy in cognitive control: Repetition expectancy modulates the congruency sequence effect. *Psychonomic Bulletin & Review*, *20*, 798–805.

Egner, T. (2008). Multiple conflict-driven control mechanisms in the human brain. *Trends in Cognitive Sciences*, *12*(10), 374-380.

Egner, T. (2014). Creatures of habit (and control): a multi-level learning perspective on the modulation of congruency effects. *Frontiers in Psychology*, *5*, 1247

Egner, T., Delano, M., and Hirsch, J. (2007). Separate conflict-specific cognitive control mechanisms in the human brain. *NeuroImage*, *35*(2007), 940-948.

Eriksen, B. A., & Eriksen, C. W. (1974). Effects of noise letters upon the identification of a target letter in a nonsearch task. *Perception & Psychophysics*, *16*(1), 143-149.

Eriksen, C. W., & Schultz, D. W. (1979). Information processing in visual search: A continuous flow conception and experimental results. *Perception & Psychophysics*, *25*, 249 –263.

Faul, F., Erdfelder, E., Lang, A.G., & Buchner, A. (2007). G*Power 3: a flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behavior Research Methods*, *39*(2), 175-191.

Fischer, R., Plessow, F., Kunde, W., & Kiesel, A. (2010). Trial-to-trial modulations of the Simon effect in conditions of attentional limitations: Evidence from dual tasks. *Journal of Experimental Psychology: Human Perception and Performance*, *36*(6), 1576-1594.

Freitas, A. L., & Clark, S. L. (2015). Generality and specificity in cognitive control: Conflict adaptation within and across selective-attention tasks but not across selective-attention and Simon tasks. *Psychological Research*, *79*, 143–162.

Frings, C., Hommel, B., Koch, I., Rothermund, K., Dignath, D., Giesen, C., ... & Philipp, A. (2020). Binding and retrieval in action control (BRAC). *Trends in Cognitive Sciences*, *24*(5), 375-387.

Folk, C. L., Leber, A. B., & Egeth, H. E. (2002). Made you blink! Contingent attentional capture produces a spatial blink. *Perception & Psychophysics*, *64*(5), 741-753.

Garner, W. R. (1974). The stimulus in information processing. *In Sensation and Measurement* (pp. 77-90). Springer, Dordrecht.

Giesbrecht, B., Woldorff, M. G., Song, A. W., & Mangun, G. R. (2003). Neural mechanisms of top-down control during spatial and feature attention. *Neuroimage, 19*, 496-512.

Giesen, C., & Rothermund, K. (2014). Distractor repetitions retrieve previous responses and previous targets: Experimental dissociations of distractor–response and distractor–target bindings. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *40*(3), 645.

Goodhew, S. C., Kendall, W., Ferber, S., & Pratt, J. (2014). Setting semantics: Conceptual set can determine the physical properties that capture attention. *Attention, Perception, & Psychophysics*, *76*(6), 1577-1589.

Grant, L. D., Cookson, S. L., & Weissman, D. H. (2020). Task Sets Serve as Boundaries for the Congruency Sequence Effect. *Journal of Experimental Psychology: Human Perception and Performance*, *46*(8), 798.

Grant, L. D., & Weissman, D. H. (2019). Turning distractors into targets increases the congruency sequence effect. *Acta Psychologica*, *192*, 31-41.

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

Hazeltine, E., Lightman, E., Schwarb, H., & Schumacher, E. H. (2011). The boundaries of sequential modulations: evidence for set-level control. *Journal of Experimental Psychology: Human Perception and Performance*, *37*(6), 1898.

Hengstler, M., Holland, R. W., van Steenbergen, H., van Knippenberg, A. (2014). The influence of approach-avoidance motivational orientation on conflict adaptation. *Cognitive, Affective, & Behavioral Neuroscience*, *14*(2), 548-560.

Hommel, B. (1998). Event files: Evidence for automatic integration of stimulus-response episodes. *Visual Cognition*, *5*(1-2), 183-216.

Hommel, B. (2004). Event files: Feature binding in and across perception and action. *Trends in Cognitive Sciences*, *8*(11), 494-500.

Hommel, B. (2007). Feature integration across perception and action: Event files affect response choice. *Psychological Research*, *71*(1), 42-63.

Hommel, B. (1997). Interactions between stimulus-stimulus congruence and stimulus-response compatibility, *Psychological Research, 59*, 248-260.

Hommel, B., Müsseler, J., Aschersleben, G., & Prinz, W. (2001). The theory of event coding (TEC): A framework for perception and action planning. *Behavioral and Brain Sciences*, *24*(5), 849.

Hommel, B., Proctor, R. W., & Vu, K. P. L. (2004). A feature-integration account of sequential effects in the Simon task. *Psychological Research*, *68*(1), 1-17.

Hommel, B., & Wiers, R. W. (2017). Towards a Unitary Approach to Human Action Control. *Trends in Cognitive Sciences*, *21*(12), 940-949.

Hsu, N. S., & Novick, J. M. (2016). Dynamic engagement of cognitive control modulates recovery from misinterpretation during real-time language processing. *Psychological Science*, *27*(4), 572-582.

Hubbard, J., Kuhns, D., Schäfer, T.A., Mayr, U. (2017). Is conflict adaptation due to active regulation or passive carry-over? Evidence from eye movements. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, *43*(3), 385-393.

Huettel, S.A., Mack, P.B., & McCarthy, G. (2002). Perceiving patterns in random series: dynamic processing of sequence in prefrontal cortex. *Nature Neuroscience*, *5*, 485-490.

Huffman, G., Hilchey, M. D., Weidler, B. J., Mills, M., & Pratt, J. (2020). Does feature-based attention play a role in the episodic retrieval of event files? *Journal of Experimental Psychology: Human Perception and Performance*, *46*(3), 241.

JASP Team (2019). JASP (Version 0.11.1)[Computer software].

Jiang, J., & Egner, T. (2014). Using Neural Pattern Classifiers to Quantify the Modularity of Conflict-Control Mechanisms in the Human Brain. *Cerebral Cortex, 24*, 1793-1805.

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

Jiménez, L., & Méndez, A. (2014). Even with time, conflict adaptation is not made of expectancies. *Frontiers in Psychology, 5*, 1042.

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

Jiménez, L., & Méndez, A. (2014). Even with time, conflict adaptation is not made of expectancies. Frontiers in Psychology, 5, 1042.

Jones, P. R. (2019). A note on detecting statistical outliers in psychophysical data. *Attention, Perception, & Psychophysics*, *81*, 1189–1196.

Kan, I. P., Teubner-Rhodes, S., Drummey, A. B., Nutile, L., Krupa, L., & Novick, J. M. (2013). To adapt or not to adapt: The question of domain-general cognitive control. *Cognition*, *129*(3), 637-651.

Kerns, J. G., Cohen, J. D., MacDonald, A. W., Cho, R. Y., Stenger, V. A., & Carter, C. S. (2004). Anterior cingulate conflict monitoring and adjustments in control. *Science*, *303* (5660), 1023-1026.

Kiesel, A., Kunde, W., & Hoffmann, J. (2006). Evidence for task-specific resolution of response conflict. *Psychonomic Bulletin & Review*, *13*(5), 800-806.

Kim, S., & Cho, Y. S. (2014). Congruency sequence effect without feature integration and contingency learning. *Acta Psychologica*, *149*, 60–68.

Kim, S., Lee, S. H., & Cho, Y. S. (2015). Control processes through the suppression of the automatic response activation triggered by task-irrelevant information in the Simon-type tasks. *Acta Psychologica*, *162*, 51-61.

Kleiman, T., Trope, Y., & Amodio, D. M. (2016). Cognitive control modulates attention to food cues: Support for the control readiness model of self-control. *Brain and Cognition*, *110*, 94-101.

Kleinsorge, T., & Heuer, H. (1999). Hierarchical switching in a multi-dimensional task space. *Psychological Research*, *62*(4), 300-312.

Koechlin, E., & Summerfield, C. (2008). The cognitive architecture of the human lateral prefrontal cortex. *Attention and Performance*, *22*.

Kornblum, S., Hasbroucq, T., & Osman, A. (1990). Dimensional overlap: cognitive basis for stimulus-response compatibility--a model and taxonomy. *Psychological Review*, *97*, 253-270.

Kreutzfeldt, M., Stephan, D. N., Willmes, K., & Koch, I. (2016). Shifts in target modality cause attentional reset: Evidence from sequential modulation of crossmodal congruency effects. *Psychonomic Bulletin & Review*, *23*(5), 1466-1473.

Kunde, W., & Wühr, P. (2006). Sequential modulations of correspondence effects across spatial dimensions and tasks. *Memory & Cognition*, *34*(2), 356-367.

Landauer, T.K. (1962). Rate of implicit speech. *Perceptual and Motor Skills*, *15*, 646.

Lavie, N. (1995). Perceptual load as a necessary condition for selective attention. *Journal of Experimental Psychology: Human Perception and Performance*, *21*(3), 451–468.

Lavie, N. (2005). Distracted and confused? Selective attention under load. *Trends in Cognitive Sciences*, *9*(2), 75–82.

Lavie, N., & Tsal, Y. (1994). Perceptual load as a major determinant of the locus of selection in visual attention. *Perception & Psychophysics*, *56*(2), 183–197.

Lee, J., & Cho, Y. S. (2013). Congruency sequence effect in cross-task context: evidence for dimension-specific modulation. *Acta Psychologica*, *144(*3), 617-627.

Lim, C. E. & Cho, Y. S. (2018). Determining the scope of control underlying the congruency sequence effect: roles of stimulus-response mapping and response mode, *Acta Psychologica*, *190*, 267-276.

Lim, C. E., & Cho, Y. S. (2020). Response mode modulates the congruency sequence effect in spatial conflict tasks: evidence from aimed-movement responses. *Psychological Research*, 1-22.

Logan, G. D. (1985). Executive control of thought and action. *Acta Psychologica*, *60*, 193–210.

Logan, G. D. (1994). On the ability to inhibit thought and action: A users' guide to the stop signal paradigm. In D. Dagenbach & T. H. Carr (Eds.), *Inhibitory processes in attention, memory, and language* (pp. 189–239). San Diego, CA: Academic Press.

Logan, G. D., & Zbrodoff, J. N. (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.

Lukas, S., Phillip, A.M., & Koch, I. (2010). The role of preparation and cue-modality in cross-modal task switching. *Acta Psychologica*, *134*, 318-322.

Martin, R., Wogalter, M., & Forlano, J. (1988). Reading Comprehension in the Presence of Unattended Speech and Music. *Journal of Memory and Language*, *27*, 382-398.

Mayr, U., Awh, E., & Laurey, P. (2003). Conflict adaptation effects in the absence of executive control. *Nature Neuroscience*, *6*, 450–452.

Mayr, S., Möller, M., & Buchner, A. (2018). Contextual modulation of prime response retrieval processes: evidence from auditory negative priming. *Attention, Perception, & Psychophysics*, *80*(8), 1918-1931.

Miles, J. D., & Proctor, R. W. (2012). Correlations between spatial compatibility effects: are arrows more like locations or words? *Psychological Research*, *76*(6), 777-791.

Miller, E. K., & Cohen, J. D. (2001). An integrative theory of prefrontal cortex function. *Annual Review of Neuroscience, 24*, 167-202.

Moore, K. S., & Weissman, D. H. (2010). Involuntary transfer of a top-down attentional set into the focus of attention: evidence from a contingent attentional capture paradigm. *Attention, Perception, & Psychophysics*, *72*, 1495–1509.

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. *Psychonomic Bulletin & Review, 19*, 750-757.

Notebaert, W., & Verguts, T. (2008). Cognitive control acts locally. *Cognition*, *106*(2), 1071-1080.

Notebaert, W., & Verguts, T. (2006). Stimulus conflict predicts conflict adaptation in a numerical flanker task. *Psychonomic Bulletin & Review*, *13*(6), 1078-1084.

Parmentier, F. B., Elsley, J. V., Andres, P., & Barcelo, F. (2011). Why are auditory novels distracting? Contrasting the roles of novelty, violation of expectation and stimulus change. *Cognition*, *119*, 374–380.

Pecher, D., Zeelenberg, R., and Barsalou, L.W. (2004). Sensorimotor simulations underlie conceptual representations: modality-specific effects of prior activation. *Psychonomic Bulletin & Review*, *11*, 164–167.

Raftery, A. E. (1995). Bayesian model selection in social research. *Sociological Methodology*, *25*, 111–163.

Ridderinkhof, K. R. (2002). Micro- and macro-adjustments of task set: activation and suppression in conflict tasks. *Psychological Research*, *66*(4), 312-323.

Robinson, T. & Berridge, K. (2008). The incentive sensitization theory of addiction: some current issues. *Philosophical Transactions of the Royal Society*, *363*, 3137–3146.

Rogers, R.D., & Monsell, S. (1995). Costs of a predictible switch between simple cognitive tasks. *Journal of Experimental Psychology: General, 124*, 615-630.

Rousseeuw, P. J., & Croux, C. (1993). Alternatives to the median absolute deviation. *Journal of the American Statistical association*, *88*(424), 1273-1283.

Scerrati, E., Baroni, G., Borghi, A. M., Galatolo, R., Lugli, L., & Nicoletti, R. (2015). The modality-switch effect: visually and aurally presented prime sentences activate our senses. *Frontiers in Psychology*, *6*, 1668.

Schmidt, J. R. (2013). Questioning conflict adaptation: proportion congruent and Gratton effects reconsidered. *Psychonomic Bulletin & Review, 20*, 615-630.

Schmidt, J. R., & Besner, D. (2008). The Stroop Effect: Why Proportion Congruent Has Nothing to Do With Congruency and Everything to Do With Contingency. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 34*, 514-523.

Schmidt, J. R., Augustinova, M., & De Houwer, J. (2018). Category learning in the color-word contingency learning paradigm. *Psychonomic Bulletin & Review, 25*, 658-666.

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*(1), 176-186.

Schmidt, J. R., & Weissman, D. H. (2014). Congruency sequence effects without feature integration or contingency learning confounds. *PLoS One*, *9*(7), e102337

Schmidt, J. R., & Weissman, D. H. (2015). Contingent attentional capture triggers the congruency sequence effect. *Acta Psychologica*, *159*, 61– 68.

Schlaghecken, F., & Maylor, E. A. (2020). When awareness gets in the way: Reactivation aversion effects resolve the generality/specificity paradox in sensorimotor interference tasks. *Journal of Experimental Psychology: General*, *149*(11).

Schumacher, E. H., & Hazeltine, E. (2016). Hierarchical Task Representation: Task files and Response Selection. *Current Directions in Psychological Science*, *25*, 449-454.

Serences, J. T., Shomstein, S., Leber, A. B., Golay, X., Egeth, H. E., & Yantis, S. (2005). Coordination of voluntary and stimulus-driven attentional control in human cortex. *Psychological Science*, *16*(2), 114-122.

Simon, J. R., & Rudell, A. P. (1967). Auditory SR compatibility: the effect of an irrelevant cue on information processing. *Journal of Applied Psychology*, *51*(3), 300.

Soto-Faraco, S., Spence, C., & Kingstone, A. (2004). Cross-modal dynamic capture: congruency effects in the perception of motion across sensory modalities. *Journal of Experimental Psychology: Human Perception and Performance*, *30*(2), 330.

Soutschek, A., Strobach, T., & Schubert, T. (2013). Working memory demands modulate cognitive control in the Stroop paradigm. *Psychological Research*, *77*(3), 333-347.

Spapé, M. M., & Hommel, B. (2008). He said, she said: Episodic retrieval induces conflict adaptation in an auditory Stroop task. *Psychonomic Bulletin & Review*, *15*(6), 1117-1121.

Spapé, M. M., & Hommel, B. (2014). Sequential modulations of the Simon effect depend on episodic retrieval. *Frontiers in Psychology*, *5*, 855.

Spence, C. (2010). Crossmodal spatial attention. *Annals of the New York Academy of Sciences*, *1191*, 182-200.

Standing, L. & Curtis, L. (1989). Subvocalization rate versus other predictors of the memory span. *Psychological Reports*, *65*, 487-495.

Stroop, J. R. (1935). Studies of interference in serial verbal reactions. *Journal of Experimental Psychology*, *18*(6), 643.

Stürmer, B., Leuthold, H., Soetens, E., Schröter, H., & Sommer, W. (2002). Control over location-based response activation in the Simon task: behavioral and electrophysiological evidence. *Journal of Experimental Psychology: Human Perception and Performance*, *28*(6), 1345-1363.

Tellinghuisen, D. J., & Nowak, E. J. (2003). The inability to ignore auditory distractors as a function of visual task perceptual load. *Perception & Psychophysics*, *65*(5), 817-828.

Turatto, M., Benso, F., Galfano, G. G., & Umiltà, C. (2002). Nonspatial Attentional Shifts Between Audition and Vision. *Journal of Experimental Psychology: Human Perception and Performance, 28*, 628-639.

Ullsperger, M., Bylsma, L. M., & Botvinick, M. M. (2005). The conflict adaptation effect: It's not just priming. *Cognitive, Affective, & Behavioral Neuroscience, 5*, 467-472.

Verbruggen, F., Notebaert, W., Liefooghe, B., & Vandierendonck, A. (2006). Stimulus- and response-conflict-induced cognitive control in the flanker task. *Psychonomic Bulletin & Review*, *13*, 328–333.

Verbruggen, F., & Logan, G. D. (2008). Automatic and controlled response inhibition: associative learning in the go/no-go and stop-signal paradigms. *Journal of Experimental Psychology: General*, *137*(4), 649.

Verguts, T., & Notebaert, W. (2008). Hebbian Learning of Cognitive Control: Dealing With Specific and Nonspecific Adaptation. *Psychological Review, 115*, 518-525.

Verguts, T., & Notebaert, W. (2009). Adaptation by binding: a learning account of cognitive control. *Trends in Cognitive Sciences, 13*, 252-257.

Wang, X., Zhao, X., Xue, G., & Chen, A. (2016). Alertness function of thalamus in conflict adaptation. *Neuroimage*, *132*, 274-282.

Weigelt, M., Güldenpenning, I., Steggemann-Weinrich, Y., Alaboud, M.A.A., & Kunde, W. (2017). Control over the processing of the opponent's gaze direction in basketball experts. *Psychonomic Bulletin & Review*, *24*(3), 828-834.

Weissman, D. H. (2020). Interacting Congruency Effects in the Hybrid Stroop-Simon Task Prevent Conclusions Regarding the Domain Specificity or Generality of the Congruency Sequence Effect. *Journal of Experimental Psychology: Learning, Memory, and Cognition.*

Weissman, D. H. (2019). Let your fingers do the walking: Finger force distinguishes competing accounts of the congruency sequence effect. *Psychonomic Bulletin & Review*, *26*(5), 1619-1626

Weissman, D. H., Colter, K., Drake, B., & Morgan, C. (2015). The congruency sequence effect transfers across different response modes. *Acta Psychologica*, *161*, 86–94.

Weissman, D. H., Colter, K. M., Grant, L.D., & Bissett, P.G. (2017). Identifying stimuli that cue multiple responses triggers the congruency sequence effect independent of response conflict. *Journal of Experimental Psychology: Human Perception and Performance*, *43*(4), 677-689.

Weissman, D. H., Egner, T., Hawks, Z., & Link, J. (2015). The congruency sequence effect emerges when the distracter precedes the target. *Acta Psychologica*, *156*, 8-21.

Weissman, D. H., Grant, L. D., & Jones, M. (2020). The congruency sequence effect in a modified prime-probe task indexes response-general control. *Journal of Experimental Psychology: Human Perception and Performance*.

Weissman, D. H., Hawks, Z. W., & Egner, T. (2016). Different levels of learning interact to shape the congruency sequence effect. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *42*, 566–583.

Weissman, D. H., Jiang, J., & Egner, T. (2014). Determinants of congruency sequence effects without learning and memory confounds. *Journal of Experimental Psychology: Human Perception and Performance*, *40*(5), 2022.

Wendt, M., Kluwe, R. H., & Peters, A. (2006). Sequential Modulations of Interference Evoked by Processing Task-Irrelevant Stimulus Features. *Journal of Experimental Psychology: Human Perception and Performance, 32*, 644-667.

Wessel, J.R., & Aron, A. R. (2017). On the Globality of Motor Suppression: Unexpected Events and Their Influence on Behavior and Cognition. *Neuron*, *93*(2), 259-280.

Wessel, J. R., & Aron, A. R. (2013). Unexpected events induce motor slowing via a brain mechanism for action-stopping with global suppressive effects. *Journal of Neuroscience*, *33*, 18481–18491.

Wessel, J. R., Jenkinson, N., Brittain, J.-S., Voets, S. H. E.M., Aziz, T. Z., & Aron, A. R. (2016). Surprise disrupts cognition via a fronto-basal ganglia suppressive mechanism. *Nature Communications*, *7*, 11195.

Wühr, P., Duthoo, W., & Notebaert, W. (2015). Generalizing attentional control across dimensions and tasks: Evidence from transfer of proportion-congruent effects. *Quarterly Journal of Experimental Psychology*, *68*(4), 779-801.

Yang, G., Nan, W., Zheng, Y., Wu, H., Li, Q., & Liu, X. (2017). Distinct cognitive control mechanisms as revealed by modality-specific conflict adaptation effects. *Journal of Experimental Psychology: Human Perception & Performance*, *43*(4), 807-818.

Yao, B., Belin, P., & Scheepers, C. (2011). Silent reading of direct versus indirect speech activates voice-selective areas in the auditory cortex. *Journal of Cognitive Neuroscience*, *23*(10), 3146-3152.

Yeung, N., Cohen, J. D., & Botvinick, M. M. (2011). Errors of interpretation and modeling: A reply to Grinband et al. *NeuroImage*, *57*, 316–319.

Zimmer, U., Roberts, K. C., Harshbarger, T. B., & Woldorff, M. G. (2010). Multisensory conflict modulates the spread of visual attention across a multisensory object. *Neuroimage*, *52*(2), 606-616.