Temporal Dynamics and Representational Consequences of the Control of Processing Conflict between Visual Working Memory and Visual Perception

Chunyue Teng[1,2], Jacqueline M. Fulvio[1], Mattia Pietrelli[1], Jiefeng Jiang[3], and Bradley R. Postle[1]

Abstract

■ Visual working memory (WM) extensively interacts with visual perception. When information between the two processes is in conflict, cognitive control can be recruited to effectively mitigate the resultant interference. The current study investigated the neural bases of the control of conflict between visual WM and visual perception. We recorded the EEG from 25 human participants (13 male) performing a dual task combining visual WM and tilt discrimination, the latter occurring during the WM delay. The congruity in orientation between the memorandum and the discriminandum was manipulated. Behavioral data were fitted to a reinforcement-learning model of cognitive control to derive trial-wise estimates of demand for preparatory and reflexive control, which were then used for EEG analyses. The level of preparatory control was associated with sustained frontal-midline theta activity preceding trial onset, as well as with the strength of the neural representation of the memorandum. Subsequently, discriminandum onset triggered a control prediction error signal that was reflected in a left frontal positivity. On trials when an incongruent discriminandum was not expected, reflexive control that scaled with the prediction error acted to suppress the neural representation of the discriminandum, producing below-baseline decoding of the discriminandum that, in turn, exerted a repulsive serial bias on WM recall on the subsequent trial. These results illustrate the flexible recruitment of two modes of control and how their dynamic interplay acts to mitigate interference between simultaneously processed perceptual and mnemonic representations. ■

Uncorrected Proof

INTRODUCTION

Holding task-relevant information in working memory (WM) while simultaneously processing sensory input is crucial for everyday tasks (e.g., Boettcher, Gresch, Nobre, & van Ede, 2021; D'Esposito & Postle, 2015; Baddeley, 1992). This mental juggling act can be challenging when the two types of information conflict (e.g., Lorenc, Mallett, & Lewis-Peacock, 2021; Bae & Luck, 2019; Teng & Kravitz, 2019; Rademaker, Bloem, De Weerd, & Sack, 2015). Although the interaction between WM and perception may be obligatory (e.g., Fukuda et al., 2022; Teng & Postle, 2021; Ding, Paffen, Naber, & Van der Stigchel, 2019; Teng & Kravitz, 2019; Kiyonaga & Egner, 2016; Rademaker et al., 2015; Gayet, Paffen, & Van der Stigchel, 2013; Soto, Hodsoll, Rotshtein, & Humphreys, 2008; Olivers, Meijer, & Theeuwes, 2006; Soto, Heinke, Humphreys, & Blanco, 2005), it is possible to mitigate interference between the two via the application of cognitive control (e.g., Teng, Fulvio, Jiang, & Postle, 2022; Kiyonaga & Egner, 2014; Kiyonaga, Egner, & Soto, 2012; Carlisle & Woodman, 2011). Here, we seek to elucidate the EEG bases of the control of conflict between WM and visual perception.

Cognitive control has been widely examined in tasks that induce conflict between stimulus–response associations, which require response inhibition, involve error processing, or impose a punishment (e.g., Egner, 2017; Aron, 2007; Botvinick, Braver, Barch, Carter, & Cohen, 2001). The dual mechanisms of the control framework (Braver, 2012) posits two distinct modes of cognitive control: proactive and reactive. Proactive control refers to the anticipatory preparation of the system so as to prioritize the processing of task-relevant information when conflict is anticipated, whereas reactive control refers to the phasic "after-the-fact" recruitment of control that is needed to resolve an unexpected instance of conflict. One way to manipulate proactive and reactive control experimentally is to do so on the time scale of blocks of trials. For example, in the color–word Stroop task (Stroop, 1935), proactive control is associated with a sustained, elevated level of attention toward task-relevant processing, as indicated by reduced average Stroop interference during blocks with a high proportion of incongruent trials (e.g., Braver, Kizhner, Tang, Freund, & Etzel, 2021; Bugg, 2012). A way to isolate the effects of reactive control, in contrast, is to

[1]University of Wisconsin–Madison, [2]Lawrence University, [3]University of Iowa

© 2025 Massachusetts Institute of Technology

Downloaded from http://direct.mit.edu/jocn/article-pdf/doi/10.1162/jocn_a_02310/2501480/jocn_a_02310.pdf?casa_token=VLDfMCJQ6UAAAAA:63PiRA1U-xfVX1sUK_IgFiFVeddBeHgHOMm8ObNCYPzfPDY8tiN688x0-HCEDB4viY_Go by University of Wisconsin, Madison, Jacqueline Fulvio on 23 April

manipulate the item-specific proportion congruency within a block. A concrete example would be to construct a block of trials such that although the words "red" and "blue" occur the same number of times during the block, instances of "red" appearing in an incongruent color are higher than those of "blue" appearing in an incongruent color. Within such a block, the cost of incongruity for a frequently incongruent item is smaller than the cost of incongruity for a frequently congruent item (cf. Gonthier et al., 2016; Bugg, 2012; Bugg & Crump, 2012). This is interpreted as evidence that the proportion-congruency manipulation affected the reactive control triggered by the two types of items because at any given point within the block, a participant cannot have predicted what would be the next word and so cannot have engaged control ahead of time. Rather, they can only do so upon presentation of the stimulus. Neuroimaging studies indicate that proactive control elicits elevated sustained activity in lateral pFC (e.g., Braver, 2012; Burgess & Braver, 2010), whereas reactive control elicits transient activation of lateral pFC (e.g., Burgess & Braver, 2010; Feredoes et al., 2006) and ACC (e.g., Chiew & Braver, 2017; Brown & Braver, 2005).

As summarized above, there is a rich theoretical framework for a dual-mode construal of cognitive control. Although that framework provides relevant context for the present work, the present work is explicitly not drawing on this framework. Rather, in this article, we will use two different words—"reflexive" and "preparatory"—in an atheoretical, descriptive manner. We will use "reflexive" as an adjective that describes a process that occurs as short-latency reaction to an unexpected event that preceded it. We will use "preparatory" as an adjective to describe a process that is carried out in advance of a future event, so as to "be ready for it to happen." Importantly, although "reflexive" implies a response that is carried out more or less instantaneously, "preparatory" does not imply a particular latency, nor a duration. For example, although saving money for retirement can be characterized as a "preparatory" activity (one that spans decades), so too can getting ready to slap when one next hears the buzz of a pesky mosquito (a behavior that plays out over seconds). Because these concepts can lend themselves to misunderstanding, we elaborate on the mosquito example with the intention of making very clear how we will use these two words. If, while reading a book, one hears a mosquito for the first time and immediately slaps in the direction of the sound, this action would be classified as "reflexive." During the ensuing few seconds, if one then sets the book in one's lap and holds perfectly still, so as to more accurately slap at the mosquito when one next hears the buzz, this newly adopted behavior would be classified as "preparatory" (i.e., by holding still, one is preparing for the anticipated return of the mosquito). At the risk of too strenuously making this point, in this anecdote, the initial buzz-plus-slap can be thought of as "trial n," the ensuing wait as an intertrial interval, and the subsequent instance of being buzzed as "trial $n+1$".

In addition to situations in which preparatory control is seen to be elevated for an extended period (e.g., across the duration of a block), an important point for the present work is preparatory control can also be deployed over relatively short time scales, as seen, for example, in trial-by-trial adjustments of the level of preparatory control (Braver, 2012). In the Stroop task, when the congruity of the previous trial is independent from the current trial, encountering incongruity between ink color and word on the previous trial (henceforth, "trial $n-1$") reduces the cost to performance of incongruity on the current trial (henceforth, "trial n"), a phenomenon known as the conflict adaptation effect (Egner, 2017; Egner & Hirsch, 2005; Weissman, Jiang, & Egner, 2014; Gratton, Coles, & Donchin, 1992). From this, one can infer that the instance of incongruity on trial $n-1$ recruits two control-related processes: First, the detection of incongruity triggers control to adjudicate the competition that exists at that moment between color naming and word naming; second, the level of control is increased in preparation for trial n. The former is considered *reflexive* control because its recruitment is in reaction to the detection of conflict—this just-detected conflict must be resolved to respond correctly on the current trial. The latter is considered *preparatory* control because it is recruited in anticipation of conflict that has not yet occurred, but that may occur in the future (i.e., on trial n). Evidence for frequent shifts of the level of preparatory control can also be seen in trial-by-trial changes in the level of activity in pFC in cued-task switching paradigms (e.g., Savine & Braver, 2010).

In the present work, we do not take a theoretical stance about whether proactive control deployed across different spans of time may arise from the same or different mechanisms. Instead, we focus on dissociating the neural signatures of preparatory and reflexive control mechanisms using a model-based approach. We recorded the EEG while participants performed a dual-task procedure that combined WM for orientation of a Gabor patch ("memorandum") with a clockwise–counterclockwise discrimination of a second Gabor patch ("discriminandum") presented during the WM delay, with the congruity of the two key experimental manipulation (Figure 1A). Although the WM and discrimination decisions were formally independent, in practice, the two interacted, a situation for which one would expect the recruitment of control. Just as the Stroop task requires controlling interference between color naming and word reading, our task requires controlling interference between representations of WM and perceptual processing. When these representations are incongruent, cognitive control is required to prevent the discriminandum from interfering with the WM representation, and to prevent the memorandum from biasing the perceptual discrimination. Previous work has shown that these control demands are behaviorally consequential: Incongruity between

Uncorrected Proof

Downloaded from http://direct.mit.edu/jocn/article-pdf/doi/10.1162/jocn_a_02310/2501480/jocn_a_02310.pdf?casa_token=VLDfMCJQ6UAAAAA:63PtRA1U-xfVX1sUK_IgFiFVeddBeHgHOMm8ObNCYPzfPDY8tiN688x0-HCEDB4wiY_Go by University of Wisconsin, Madison, Jacqueline Fulvio on 23 April

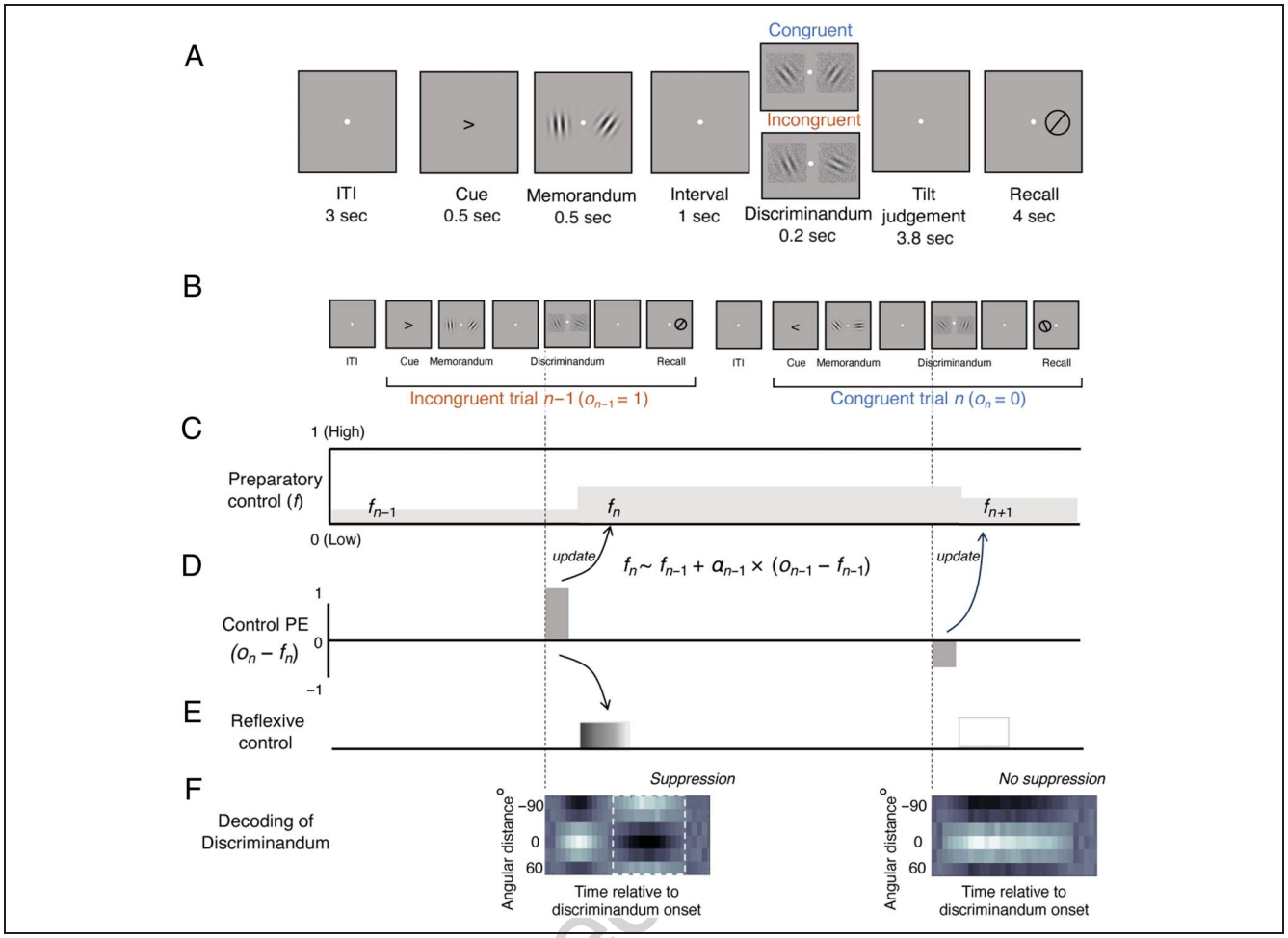


Figure 1. Task structure and the hypothesized interplay between preparatory and reflexive control. (A) Each trial started with a centrally presented arrow cue that indicated which side of the screen to attend to for both memory and discrimination tasks. Participants first memorized the orientation of a Gabor patch ("memorandum"), and then during the WM delay, they performed a speeded clockwise/counterclockwise tilt judgment on the noise-covered Gabor patch ("discriminandum"). The orientation of the discriminandum was congruent (i.e., identical) with that of the memorandum on half of the trials, and it was incongruent (i.e., differing by ±30°, ±60°, or 90°) with it on the other half of the trials. (B) Illustration of two consecutive trials of the dual WM + discrimination task, with the orientation of the discriminandum incongruent with the memorandum on the preceding trial illustrated ("*n*–1") and congruent on current trial *n*. (C) The dynamics of preparatory control across the two trials illustrated in B. In the FCM, preparatory control is operationalized by the *predicted conflict* (f) parameter, and here, it is illustrated with gray shading. For the situation illustrated here, the value of f preceding the onset of trial *n*–1 is relatively low, and this persists at this level until it is updated by the *control PE* (D) that is itself triggered by the onset of that trial's discriminandum. Because the discriminandum on trial *n*–1 is incongruent with the memorandum, the value of f increases, and this elevated level of preparatory control persists until the onset of the discriminandum on trial *n*. (D) The *control PE* on these two trials. On each trial, upon discriminandum onset, a signed *control PE* signals the discrepancy between the current level of f and the observed congruity (o_n- f_n). On trial *n*–1, incongruity results in a positive *control PE* (because o_{n-1} takes a value of 1), whereas on trial *n*, congruity results in a negative *control PE* (because o_n takes a value of 0). The *control PE* has two functions. First, it is used to update f. Second, it triggers phasic reflexive control. (E) The relatively large positive *control PE* on trial *n*–1 triggers a relatively strong recruitment of reflexive control, whereas the negative *control PE* on trial *n* results in no recruitment of reflexive control (note that the effects of the *control PE* are assumed to be task specific: In this task, an incongruent discriminandum recruits reflexive suppression [so as to minimize interference with recall of the memorandum], but a congruent discriminandum does not require the application of control). (F) Hypothesized decoding results of the discriminandum, plotted as pattern similarity between the test orientation and representational templates from the training set. Note that, following the discrimination response, a representation of the discriminandum is no longer needed. On trial *n*–1, reflexive control is implemented as suppression of the neural representation of the incongruent discriminandum. On trial *n*, in the absence of reflexive control, the representation of the discriminandum gradually decays. See Figure 1-1 for a visual representation of trial-by-trial dynamics of parameter fits for 30 consecutive trials from a representative participant.

memorandum and discriminandum (i.e., different orientations) impairs both visual discrimination performance and WM recall precision (Teng et al., 2022). Furthermore, fits of these data to a reinforcement learning-based flexible control model (FCM; Jiang, Beck, Heller, & Egner, 2015; Jiang, Heller, & Egner, 2014) illustrate the adaptive adjustment of control: The level of predicted conflict is negatively related to the behavioral cost of memorandum-discriminandum incongruity; unexpected incongruity triggers reflexive suppression of the incongruent discriminandum, such that it exerts a repulsive serial bias on WM recall on the subsequent trial, and it impairs

Uncorrected

Downloaded from http://direct.mit.edu/jocn/article-pdf/doi/10.1162/jocn_a_02310/2501480/jocn_a_02310.pdf?casa_token=VLDfFMCJQ6UAAAAA:63PtRA1U-xfVX1sUK_IgFiFVeddBeHgHOMm8ObNCYPzfPDY8tiN688x0-HCEDB4viY_Go by University of Wisconsin, Madison, Jacqueline Fulvio on 23 April

subsequent processing around its featural value (Teng et al., 2022). (Congruent discriminanda, in contrast, exert the more commonly observed attractive serial bias, and they enhance processing around its featural value.) In this framework, we consider the effects of control to be "reflexive" or "preparatory" based on their timing relative to the event that triggered the recruitment of control (i.e., to the onset of the discriminandum on trial n–1). Specifically, when the discriminandum on trial n–1 is incongruent (i.e., when it does not match the memorandum from trial $n-1$), reflexive control is defined as the consequent differential processing of that discriminandum. The effects of this reflexive control can be measured in the RT and accuracy of the discrimination response, the ERP that is time-locked to the discriminandum, and in the bias exerted by the discriminandum from trial n–1 on performance on trial n. Preparatory control, in contrast, is operationally defined as the anticipatory adjustments in control that influence EEG activity during the intertrial interval (ITI) following trial n–1, as well as in behavioral and EEG/ERP correlates of the processing of the memorandum on trial n. Our goal with the present work was to use fits of the FCM to behavioral data from the WM + discrimination dual task (Teng et al., 2022) to identify and distinguish EEG correlates of preparatory and reflexive control processes recruited to control WM–perception interactions.

METHODS

Participants

Twenty-five participants (12 female and 13 male, mean age = 22.3 ± 4.6) were recruited from the community of the University of Wisconsin–Madison in exchange for monetary compensation. The sample size was determined based on previous studies that investigated similar ERP signatures of frontal midline theta power in the context of cognitive control (e.g., Messel, Raud, Hoff, Stubberud, & Huster, 2021; Whitehead, Ooi, Egner, & Woldorff, 2019; Töllner et al., 2017; Cohen, 2016). All participants reported being neurologically healthy, right-handed, and having normal or corrected-to-normal visual acuity. The study was approved by the University of Wisconsin–Madison Health Sciences Institutional Review Board.

Stimuli and Procedure

The stimuli were created and presented using MATLAB (The Mathworks) and the Psychtoolbox extension (Brainard, 1997). They were displayed on a 60 × 34 cm screen with a refresh rate of 60 Hz. Participants viewed the stimuli from a distance of approximately 58 cm in a dimly lit room.

The experiment had a dual-task structure: WM delayed recall (a.k.a. "delayed estimation") with an independent perceptual discrimination occurring during the WM delay period. The background was gray throughout the experiment. Each trial began with a centrally presented left- or right-pointing arrow (black, 0.8° × 0.8°), its direction indicating the side of screen to attend to for that trial. After 0.5 sec, two memoranda (Gabor patches: radius = 6°; contrast = 0.8; spatial frequency = 4 cycles/°; phase angle randomized between 0° and 180°; 10° apart) were presented, centered on the horizontal meridian, one on each side of a white central fixation dot, for 0.5 sec. Participants were instructed to memorize the orientation of the memorandum that appeared on the cued side of the screen, but to maintain central fixation. After a delay of 1 sec, two discriminanda (Gabor patches, same properties as the memoranda) appeared at the same two locations for 0.2 sec. Participants were instructed to judge the tilt of the discriminandum on the cued side relative to vertical and indicate their response by pressing, with fingers of the left hand, the "E" key for a clockwise tilt and the "F" key for a counterclockwise tilt. After the offset of the discriminanda for 2.8 sec, a recall dial, displaying a randomly determined orientation, appeared at the location of the cued memorandum, and participants were instructed to press, with fingers of the right hand, the left and right arrow keys to rotate the dial so as to replicate their memory of the orientation of the memorandum. The probe response window was 4 sec. An ITI of 3 sec separated each trial.

The direction of the cue was randomly determined on each trial. The orientation of the memorandum was independently selected, with replacement, from a fixed set of six values spaced by 30° (15°, 45°, 75°, 105°, 135°, and 165°) with a jitter between 0° and 3° added. The congruity in orientation between memorandum and discriminandum was manipulated and counterbalanced across trials. On half of the trials, the two orientations matched, whereas on the other half, they were either ±30°, ±60°, or 90° apart. This manipulation resulted in six orientation values for the discriminandum, 15°, 45°, 75°, 105°, 135°, and 165°, the same as the memorandum values. The memorandum had a fixed contrast of 0.8. To enhance task engagement, the contrast of the discriminandum was varied and counterbalanced among five levels, (0.09, 0.07, 0.32, and 0.6). Noise was added to the discriminanda by replacing 30% of their pixels with random luminance. The orientations of the items in the uncued visual fields were also independently selected from the same pool of orientations. Twenty-four participants completed eight 60-trial blocks, and one participant completed seven 60-trial blocks. Participants completed the experiment in one session.

Behavioral Analysis

Performance on the discrimination task was assessed via RT and accuracy. Responses that were incorrect, or 3 *SD*s above or below the mean RT, were excluded from the RT analyses (approximately 3%). Note that, although we assume that incongruity on trial n–1 triggers reflexive control of the processing of that trial's incongruent discriminandum (Figure 1E), there is not clear evidence of the effects of this control on discrimination performance on that trial. Instead, evidence that the processing of the

Uncorrected Proof

Downloaded from http://direct.mit.edu/jocn/article-pdf/doi/10.1162/jocn_a_02310/2501480/jocn_a_02310.pdf?casa_token=VLDfMCJQ6UAAAAA:63PtRA1U-xfVX1sUK_IgFFVeddBeHgHOMm8ObNCYPzfPDY8tiN688x0-HCEDB4wiY_Go by University of Wisconsin, Madison, Jacqueline Fulvio on 23 April 2024

discriminandum on trial n–1 was influenced by reflexive control can be seen in the bias that it exerts on WM recall on trial n (Teng et al., 2022). We examined the influence of congruity of trial n–1 on RT and accuracy of WM recall on trial n. This analysis builds on a long history of past research on conflict adaptation and sequential trial effects that cognitive control can adapt dynamically based on recent experiences of conflict (e.g., Braem et al., 2019; Egner, 2017; Gratton et al., 1992).

WM performance was measured in terms of recall bias and precision. Both within-trial and cross-trial effects were analyzed.

Within-trial Effects

Bias. To assess within-trial recall bias, we computed the mean signed error for each of the distance conditions between the trial n memorandum and the trial n discriminandum. A positive value would indicate an attractive bias of trial n recall toward trial n's discriminandum, whereas a negative value would indicate a repulsive bias.

Precision. Within-trial precision was determined by sorting trials based on the distance between the memorandum and discriminandum on the current trial n, and the inverse of the mean absolute error (precision = 1/[mean absolute error]) was calculated for each distance condition.

Cross-trial Effects

Cross-trial serial bias. We used serial bias to assess the residual influence of the discriminandum under different conditions (congruent vs. incongruent; whether its congruity is consistent with predicted conflict or not). It is known that visual perception and visual memory of a current stimulus are influenced by previous stimuli, with the direction of the bias being either attractive or repulsive. Attractive serial bias has been observed when information in WM decays passively (e.g., Shan & Postle, 2022) and is often interpreted as the influence of previous trial decisions on a postperceptual Bayesian decision process (e.g., Bliss, Sun, & D'Esposito, 2017; Fischer & Whitney, 2014). And because sensory adaptation, in contrast, is thought to generate repulsive biases (Pascucci et al., 2019), the sign of the serial bias observed on a given trial is believed to reflect which of these factors—adaptation-related repulsion versus decision-related attraction—is exerting more influence. For example, the fact that attractive serial bias is observed under conditions that encourage passive decay from WM, but repulsive serial bias is observed under conditions that encourage active removal of information from WM (Shan & Postle, 2022), suggests a link between repulsive bias and flexible control.

It is noteworthy that Teng and colleagues (2022) reported repulsive bias for incongruent discriminanda only when their incongruity was unexpected, and thus, the conflict could only be resolved reactively. This bias was not observed when the incongruity was expected and thus could be managed with preparatory control. We have speculated that reflexive control in this task context may involve the active recruitment of perceptual adaptation mechanisms so as to suppress the processing of potentially interfering stimuli. In the present study, using a task design similar to Teng and colleagues (2022), we aimed to replicate the behavioral findings and test the prediction that the repulsive serial bias observed with unexpected incongruent discriminanda will be associated with the suppression of their neural representations as measured with EEG.

We performed cross-trial serial dependence analyses with both model-free and model-based approaches. For the model-free approach, we calculated the median error (signed) for trials for which the relative distance between the discriminandum from trial n–1 and the memorandum from trial n ranged from 0 to ±45°, and subtracted that from the median error on all trials within that range (e.g., Teng et al., 2022; Samaha, Switzky, & Postle, 2019).

For the model-based analysis of serial dependence, we employed the derivative of Gaussians procedure (Teng et al., 2022; Samaha et al., 2019; Bliss et al., 2017; Fritsche, Mostert, & de Lange, 2017): First, to prepare the data, trials with absolute error larger than 45° were excluded (to minimize influences from random guesses). Next, we demeaned the data by subtracting the mean signed error across all trials from each trial's error for each participant (to remove systematic response bias).

Subsequently, trials were sorted by the relative distance between trial n–1 discriminandum and trial n memorandum. The data were smoothed by a 15-trial moving median filter, following previous conventions (e.g., Teng et al., 2022; Samaha et al., 2019). We then fit a derivative-of-Gaussian (DoG) function (Bliss et al., 2017; Fritsche et al., 2017; Fischer & Whitney, 2014) to the data:

$$y = xawce^{-(wx)^2}$$

in which x is the relative orientation of trial n–1 discriminandum, a corresponds to the amplitude of the peaks of the curve, w corresponds to the width of the curve, and c is a constant.

The DoG was fitted to the group-averaged data with two free parameters, amplitude (a; allowed to vary between −15° and 15°) and width (w; allowed to vary between 0.02 and 0.2). A positive amplitude indicates an attractive serial bias; a negative amplitude indicates a repulsive serial bias. Width (w) scales the width of the tuning. The statistical significance of the DoG fits was determined with a bootstrapping procedure (10,000 interactions) in which we sampled 25 participants with replacement and fit the DoG function to the group average data for each interaction. This procedure resulted in 10,000 values of the amplitude of the fit, and we compared this distribution against 0 to obtain the two-tailed p value.

Uncorrected Proof

Downloaded from http://direct.mit.edu/jocn/article-pdf/doi/10.1162/jocn_a_02310/2501480/jocn_a_02310.pdf?casa_token=VLDfMCJQ6UAAAAA:63PiRA1U-xfVX1sUK_IgFiFVeddBeHgHOMm8ObNCYPzfPDY8tiN688x0-HCEDB4wiY_Go by University of Wisconsin, Madison, Jacqueline Fulvio on 23 April

Cross-trial precision. To examine the influence of the trial n–1 discriminandum on memory recall precision on trial n, trials were sorted based on the distance between trial n–1 discriminandum and trial n memorandum, and precision calculated as the inverse of the mean absolute error for each distance bin.

EEG Acquisition and Preprocessing

Sixty-three channels of EEG were recorded with BrainVision Recorder using active Ag/AgCl electrodes feeding into an ActiCHamp amplifier (Brainprodcuts GmbH). The electrodes were placed in accordance with the extended International 10–20 system. Impedances were kept below 25 kΩ. The data were sampled at 1000 Hz. An electrode placed on the forehead served as ground. All channels were referenced online to channel Fpz and rereferenced offline to the average of all electrodes.

Preprocessing and analysis of the data were completed in MATLAB (2021, Mathworks Inc.) using custom routines and functions from EEGLab (Delorme & Makeig, 2004) and Fieldtrip (Oostenveld, Fries, Maris, & Schoffelen, 2011). The data were downsampled to 500 Hz and filtered offline with a band-pass filter (0.1 Hz–100 Hz). Bad channels were identified by visual inspection and replaced via the spherical spline interpolation method. The data were epoched around the onset of the cue (−1000 to 1000 msec) and the onset of the discriminandum (−500 to 1000 msec). Epochs with amplitudes above a threshold of ± 100 μV were excluded from further analyses (M = 3.9%, SD = 3.17%). Artifacts related to eye blinks, horizontal eye movements, and muscles were corrected using an independent component analysis (M = 2.6 components, SD = 1.1 components).

ERP

For ERP analyses, an additional low-pass filter (30 Hz, second-order Butterworth) was applied to the preprocessed data. The time window of interest was determined to be −200 to 800 msec around the onset of the discriminandum, and the data were baseline corrected from −200 to 0 msec.

Time–Frequency Analysis

The preprocessed data were decomposed with a complex Morlet wavelet (1–35 Hz, 1-Hz frequency bins, 3–10 cycles; Cohen, 2019) with a custom routine in the Fieldtrip toolbox (Oostenveld, Fries, Maris, & Schoffelen, 2011). We then downsampled the time–frequency results to 100 Hz to increase processing efficiency for subsequent analyses. The data were baseline corrected to the average of all trial ITI in a condition-nonspecific way from the window from −1000 to −800 msec relative to cue onset with the decibel conversion. As a result, the baseline correction would not affect the results of subsequent analyses, regardless of the specific period chosen. (Note that, because of the anticipatory nature of preparatory control, the time window of interest began before the trial onset, i.e., at −800 msec.)

FCM

The levels of trial-wise preparatory and reflexive control were quantified with the FCM (Jiang et al., 2014, 2015). The FCM uses a reinforcement learning algorithm with a flexible learning rate to account for the flexible adjustment of control in a changing environment and provides trial-by-trial estimates of level of demand for preparatory and for reflexive control. It has been successfully applied to conflict adaptation tasks, such as a modified Stroop task (Jiang et al., 2014), revealing an insula-frontostriatal network involved in predicting and executing control (Jiang et al., 2015). It has also been used to elucidate the role of lateral pFC in anticipatory control, via its application to data from an experiment applying TMS to this region (Muhle-Karbe, Jiang, & Egner, 2018). Most recently, we have used the FCM to investigate WM–perception interactions on a behavioral task similar to the one presented here (Teng et al., 2022).

Details of the model have been described previously in Jiang and colleagues (2014, 2015), and its application a concurrent WM–perception tasks has been outlined in Teng and colleagues (2022). In summary, the model consists of three key parameters: predicted level of conflict (f), observed trial congruity (o), and a flexible learning rate (α). The observed trial congruity (o) is a binary variable (0 = congruent, 1 = incongruent) serving as the input to the model, whereas the model provides trial-wise estimates for the other two parameters. The model maintains a joint probabilistic distribution of the predicted conflict level (f) and the flexible learning rate (α). f Represents the model's anticipation of the probability of the forthcoming trial being incongruent (continuous, $f \in [0, 1]$), thereby operationalizing the level of demand for preparatory control. The flexible learning rate (α) corresponds to the model's belief in the volatility in the environment and was not included in our analyses because volatility was not experimentally manipulated, and thus not within the scope of our study, nor did it directly relate to our theoretical framework. For each trial n, the model has made a prediction of the level of conflict that will be present on that trial (f_n) based on the joint distribution of f and α. As the trial gets underway, the congruity between memorandum and discriminandum (o_n) is observed, and the distribution is updated based on a reinforcement learning rule: $f_{n+1} \leftarrow f_n + \alpha \times (o_n - f_n)$. Importantly, the "control prediction error" (*control PE*) corresponding to the difference between a trial's predicted conflict f_n and its subsequently observed congruity o_n (i.e., $o_n - f_n$) has two roles: quantifying the amount of reflexive control required by the discriminandum (thereby triggering the recruitment of reflexive control) and updating f_n (for the remainder of trial$_n$ and persisting across the ITI and into trial$_{n+1}$ until

Uncorrected Proof

Downloaded from http://direct.mit.edu/jocn/article-pdf/doi/10.1162/jocn_a_02310/2501480/jocn_a_02310.pdf?casa_token=VLDfMCJQ6UAAAAA:63PtRA1U-xfVX1sUK_IgFiFVeddBeHgHOMm8ObNCYPzfPDY8tiN688x0-HCEDB4wiY_Go by University of Wisconsin, Madison, Jacqueline Fulvio on 23 April

it is updated by trial$_{n+1}$'s discriminandum). On incongruent trials, a positive *control PE* is generated (between 0 and 1), and on congruent trials, a negative *control PE* is generated (between −1 and 0). For analyses that compared tertiles of trials of highest and lowest *control PE*, trials were selected separately for congruent and incongruent conditions. For incongruent trials, the tertile of trials with the highest positive *control PE* values (largest in absolute value) were selected, and for congruent trials, the tertile of trials with the lowest negative *control PE* values (largest in absolute value) were selected.

Figure A1 provides a visual representation of trial-by-trial fluctuation of parameter fits for 30 consecutive trials from a representative participant. In addition, to demonstrate that *f* integrates both recent ($n-1$) and distant (before n–1) trial history, we performed a regression analysis on *f* with regressors of trial congruity from trials $n-3$, $n-2$, and n–1. The regression weights, shown in Figure A2, indicate that *f* is influenced by trial congruity at least three trials back, consistent with a preparatory control mechanism that integrates over recent experiences to optimize performance.

We performed a general linear model analyses across time and space to investigate the neural correlates of preparatory and reflexive control in the EEG signals. For reflexive control, because its implementation is hypothesized to be effectively instantaneous, we reasoned that the corresponding immediate, phasic responses evoked by unexpected stimuli would be best captured through ERPs, which offer finer-grained temporal resolution than do spectral measures. To capture the stimulus-evoked effects of the *control PE*, for each participant, we performed a multiple regression on the voltage data for all time points of interest within a window from −200 to 800 msec around discriminandum onset at each electrode. Regressors included *f*, magnitude of *control PE* (absolute value of *control PE*), and congruity.

In contrast, shifts of preparatory control were expected to persist through trials, likely to span from one trial to the next (e.g., the level of preparatory control adjusted as a result of an event during trial n will, by definition, persist during the ITI and into the following trial.) Such extended levels of "state" are better captured by spectral analyses. To characterize the hypothesized sustained effect of preparatory control, we first calculated narrowband theta-, alpha-, and beta-band power by averaging across the frequency intervals of 4–7 Hz, 8–13 Hz, and 14–30 Hz, respectively. Then, we performed the same type of regression analysis on the power of each frequency band for each time point within a window from −1000 to 1000 msec around trial onset at each electrode, with *f* as the regressor of interest (combined with a constant term to form a general linear model).

For the effects of reflexive control in the discriminandum-locked ERP, significance testing was carried out on the regression weights of *f, control PE*, and congruity with comparisons against 0. For the effects of preparatory control in theta power, we compared the regression weights of *f* against 0. Multiple comparisons were corrected for with a nonparametric cluster-based permutation test across time points and channels (10,000 permutations with the Monte Carlo method; Maris & Oostenveld, 2007). First, adjacent data points (one data point apart) and channels (a minimum of two required) with a p value smaller than .05 (by two-tailed t test) were clustered together. Then, we used the maxsum method (the sum of the t values within each cluster) to determine the cluster-level statistic and comparing it to the maxsum of clusters identified in the null distribution obtained from the 10,000 permutations. For the completeness of the report, the same regression analysis was performed on the averaged band power of theta, alpha, and beta frequencies during the discriminandum phase.

Multivariate Decoding

Time-resolved multivariate decoding of the orientations of the memorandum and the discriminandum was performed on the voltage data of 17 posterior channels (P7, P5, P3, P1, Pz, P4, P6, P8, PO7, PO3, POz, PO4, PO8, O1, Oz, and O2), following Wolff, Jochim, Akyürek, and Stokes (2017) and Wolff, Ding, Myers, and Stokes (2015). The preprocessed data were downsampled to 50 Hz to increase computational efficiency. We decoded orientation using Mahalanobis distance to compute the distances between the orientation on a given trial and the full range of possible orientations. Decoding was performed separately and independently for memorandum and discriminandum, and for items presented to the left and right of fixation within each participant. We used a leave-one-trial-out cross-validation method to calculate the decodability of the trial-wise orientation. For each time point of a given trial, its activity pattern was compared with each of the six orientation bins formed with the remainder of trials (−75°, −45°, −15°, 15°, 45°, 75°). The Mahalanobis distance between the left-out trial and each orientation bin was calculated with the covariance matrix from the training trials using a shrinkage estimator. The number of trials for each orientation bin was kept equal by randomly subsampling to match the smallest number of condition-specific trials. This procedure was repeated 100 times to obtain reliable estimates.

For the visualization of tuning curves, the pairwise distances were mean-centered and the sign was reversed. Thus, higher values indicate greater pattern similarity between a pair of orientations and lower values indicate less similarity between a pair. Decoding accuracy was calculated as the slope of the tuning curves. Lastly, for significance testing, cluster-based signed permutation tests were done in windows anchored to memorandum onset (0–800 msec) and to discriminandum onset (0–800 msec), with a cluster-forming and cluster-testing significance threshold of 0.05. Time-resolved decoding values were

Uncorrected Proof

Downloaded from http://direct.mit.edu/jocn/article-pdf/doi/10.1162/jocn_a_02310/2501480/jocn_a_02310.pdf?casa_token=VLDfMCJQ6UAAAAA:63PtRA1U-xfVX1sUK_IgFiFVeddBeHgHOMm8ObNCYPzfPDY8tiN688x0-HCEDB4viY_Go by University of Wisconsin, Madison, Jacqueline Fulvio on 23 April

smoothed with a Gaussian kernel (3 time points) for visualization and statistical tests.

The decoding results were used to test two predictions about the effects of control on stimulus representation. First, to assess the possible influence of preparatory control on enhancing the processing of the memorandum, we tested whether orientation decoding of the memorandum was higher for the tertile of trials with the highest *f* than the tertile of trials with the lowest *f* (permutation procedure; 10,000 permutations).

Second, to test the prediction that incongruent discriminanda with high *control PE*s would elicit reflexive suppression, we compared the averaged decoding accuracy of the tertile of incongruent trials with the highest *control PE* versus the tertile of incongruent trials with the lowest *control PE* (each against zero, between the two; permutation procedure; 10,000 permutations with cluster correction). If unexpectedly incongruent discriminanda are indeed suppressed, their neural representation should be markedly different from those of the same stimuli on trials when they serve as congruent discriminanda. To examine this, we first constructed the representational profile for the memorandum by training a decoding model on the orientations of the memorandum. We then applied this model to decode the orientations of the discriminandum. During the testing phase, each test trial was labeled by the orientation of the discriminandum and we computed the neural distance between this test trial and the memorandum profiles. If a discriminandum and a memorandum share the same neural code, a discriminandum with a specific orientation (D1) would exhibit the closest distance to the memorandum of the same orientation (M1) relative to all other orientations (M2–M6) in the memorandum space. Alternatively, if D1 is suppressed, its neural distance to M1 would be the furthest relative to all other stimuli M2–M6. Therefore, we predicted that incongruent discriminanda generating a high *control PE*s would reconstruct an inversed similarity pattern for that required additional involvement of reflexive control. Similar approaches have been reported in previous studies (Zhang & Luo, 2023; Yu, Teng, & Postle, 2020).

RESULTS

Within-trial Behavioral Results

Trials were sorted by the distance in orientation (0°, ±30°, ±60°, and 90°) between the memorandum and the discriminandum within the same trial n. For memory precision, a repeated-measures ANOVA revealed a main effect of orientation distance on memory precision, $F(3, 72) = 12.2$, $p < .001$. Further pairwise comparisons showed elevated precision in the congruent condition (0° distance) relative to all three incongruent distances (±30°, ±60°, and 90°; ps $< .001$, Bonferroni-corrected) and no differences among the incongruent distances (all ps $> .17$). For bias in memory recall, the main effect of orientation distance was significant, $F(3, 72) = 3.53$, $p = .019$, with greater attractive bias in ±30°condition than in the congruent condition (0° distance; $p = .02$, Bonferroni-corrected).

Cross-trial Behavioral Results

Model-free Analyses, Visual Discrimination

We first performed model-free analyses to investigate whether the congruity of trial n–1 influenced visual discrimination performance on trial n. The rationale was to isolate preparatory control by examining how previous trials influence current performance, reflecting ongoing adjustments in cognitive strategy. Repeated-measures ANOVAs with factors of trial n–1 congruity and trial n congruity were conducted on discrimination RT and discrimination accuracy. For discrimination RT, we observed a significant main effect of trial n–1 congruity, $F(1, 24) = 4.68$, $p = .041$, indicating that RT on trial n was faster following a congruent trial n–1 compared with an incongruent trial n–1. Neither the main effect of trial n congruity nor the trial n–1 × trial n interaction reached significance Fs < 0.26, ps $> .61$ (Figure 2A). In terms of discrimination accuracy, the main effect of trial n congruity was significant, $F(1, 24) = 8.00$, $p = .01$, whereas the effect of trial n–1 congruity was not, $F(1, 24) < 0.001$, $p = .99$. Furthermore, the congruity effect on trial n (the difference between congruent and incongruent condition) was reduced if trial n–1 was incongruent compared with when it was incongruent, as indicated by a significant Trial n–1 × Trial n interaction, $F(1, 24) = 11.15$, $p = .003$, a Gratton effect (Figure 2A).

Model-free Analyses, WM Recall

Previous work with the dual-task procedure featured here (Teng et al., 2022) and with a WM retrocuing task (Shan & Postle, 2022) has provided evidence that when a stimulus elicits vigorous reflexive control (e.g., an unexpected incongruent discriminandum), it can exert a repulsive serial bias, a reversal of sign of its influence when reflexive control is not needed. To assess evidence for this phenomenon in the present experiment, trials were sorted by congruity, and a Difference of Gaussians (DoG) fitting of the bias was applied. Results indicated that although congruent discriminanda on trial n–1 exerted the expected attractive bias on WM recall on trial n ($a = 1.88°$, $p = .002$, 95% CI [0.94°, 2.70°], bootstrapped), incongruent discriminanda on trial n–1 did not have a significant influence on WM recall on trial n ($a = -0.61°$, $p = .10$, 95% CI [−1.45°, 0.18°], bootstrapped). The amplitudes of the two curves were significantly different from each other ($p < .001$, bootstrapped; Figure 3A). (Note that we believe that the oscillatory patterns observable in Figure 3A are due to noise because there is no systematic variation in trial structure [i.e., memorandum orientations and distractor congruity were randomly assigned from trial to trial]. Notably, this pattern is absent in Figure 3B, which illustrates the

Uncorrected Proof

Downloaded from http://direct.mit.edu/jocn/article-pdf/doi/10.1162/jocn_a_02310/2501480/jocn_a_02310.pdf?casa_token=VLDfMCJQ6UAAAAA:63PiRA1U-xfVX1sUK_IgFiFVeddBeHgHOMm8ObNCYPzfPDY8tiN688x0-HCEDB4viY_Go by University of Wisconsin, Madison, Jacqueline Fulvio on 23 April

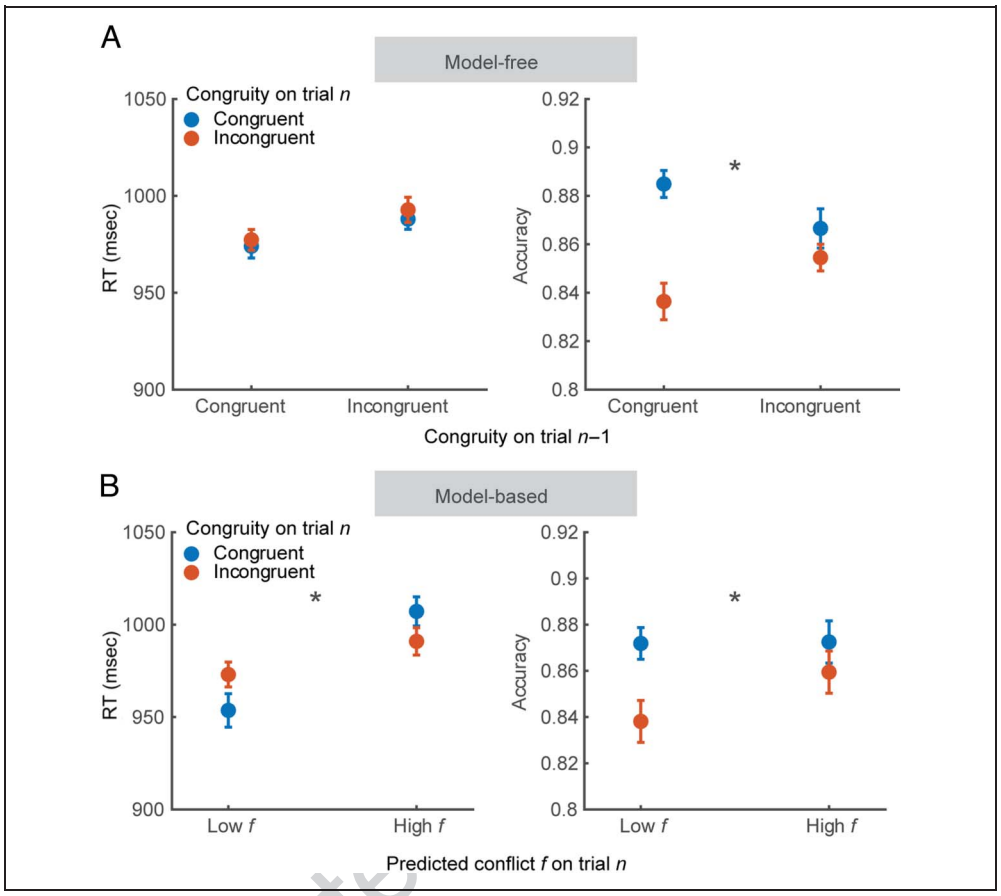


Figure 2. Behavioral effects of conflict adaptation on visual discrimination. (A) RT and accuracy on the discrimination task on trial n, as a function of trial n–1 and trial n congruity. For accuracy, the congruity effect on trial n was reduced following an incongruent trial n–1: a reduction in both the interference caused by an incongruent discriminandum and the benefit caused by a congruent one, compared with following a congruent trial n–1. (B) RT and accuracy replotted based on FCM estimates of predicted conflict f on trial n. f Modulated the congruity effects for both RT and accuracy. Figure A2 demonstrates that the parameter estimate of predicted conflict f captures a broad range of task history. * Indicates $p < .05$ for the interactions; error bars indicate standard error (*SE*).

same data sorted in a different way and which is a direct test our theoretical hypothesis.) Further analysis showed no significant differences in serial bias across differing levels of incongruity (±30°, ±60°, and 90° difference), validating the aggregation of those trials in subsequent analyses. Note that for the attractive bias observed after a congruent trial, one cannot disentangle the potential influence of the memorandum from that of the discriminandum. We further assessed the serial bias of the previous trial's memorandum on incongruent trials, which revealed an attractive bias toward the previous trial's memorandum on incongruent trials ($a = 2.02°$, $p < .001$, 95% CI [1.01°, 3.12°], bootstrapped).

Finally, precision in WM recall on trial n was influenced by trial n congruity, $F(1, 24) = 11.56$, $p < .001$, but not by trial n–1 congruity, $F(1, 24) = 3.78$, $p = .064$, nor was the interaction significant, $F(1, 24) = 1.09$, $p = .31$.

Model-based Assessment, Preparatory Control

We operationalized the level of preparatory control using the f parameter (predicted conflict) from the FCM, which was found to modulate performance in visual discrimination. Trials were sorted by f, and the tertile of trial with the highest f and the tertile of trials with the lowest f were selected for further analyses. First, repeated-measures ANOVAs were conducted on discrimination RT and discrimination accuracy on trial n, with factors of f from trial n and congruity on trial n. (This has the effect of assessing the level of preparatory control that was present at the onset of trial n.) For discrimination RT, we observed a significant main effect of f, $F(1, 24) = 9.42$, $p = .005$, and a significant interaction between f and congruity, $F(1, 24) = 7.24$, $p = .013$. Specifically, trials that began with a higher f exhibited a reduced congruity effect (Figure 2B). For discrimination accuracy, we found a main effect of congruity, $F(1, 24) = 7.61$, $p = .011$, and a significant $f \times$ Congruity interaction, $F(1, 24) = 4.36$, $p = .048$, reflecting the fact that the congruity effect was absent in the presence of high f (Figure 2B). For WM recall precision, only a significant main effect of trial n congruity was observed, $F(1, 24) = 46.20$, $p < .001$ (for the other effects, Fs < 0.60, ps > .41).

Model-based Assessment, Reflexive Control

The model-free analyses indicated that, although the delayed recall and visual discrimination elements of the dual task are formally independent, they nonetheless interact, in that incongruity between memorandum and discriminandum worsens discrimination performance and WM recall precision (replicating Teng et al., 2022). It follows from this that the cognitive system may recruit reflexive control in instances in which, when expected

Uncorrected

Downloaded from http://direct.mit.edu/jocn/article-pdf/doi/10.1162/jocn_a_02310/2501480/jocn_a_02310.pdf?casa_token=VLDfFMCJQ6UAAAAA:63PiRA1U-xfVX1sUK_IgFiFVeddBeHgHOMm8ObNCYPzfPDY8tiN688x0-HCEDB4viY_Go by University of Wisconsin, Madison, Jacqueline Fulvio on 23 April

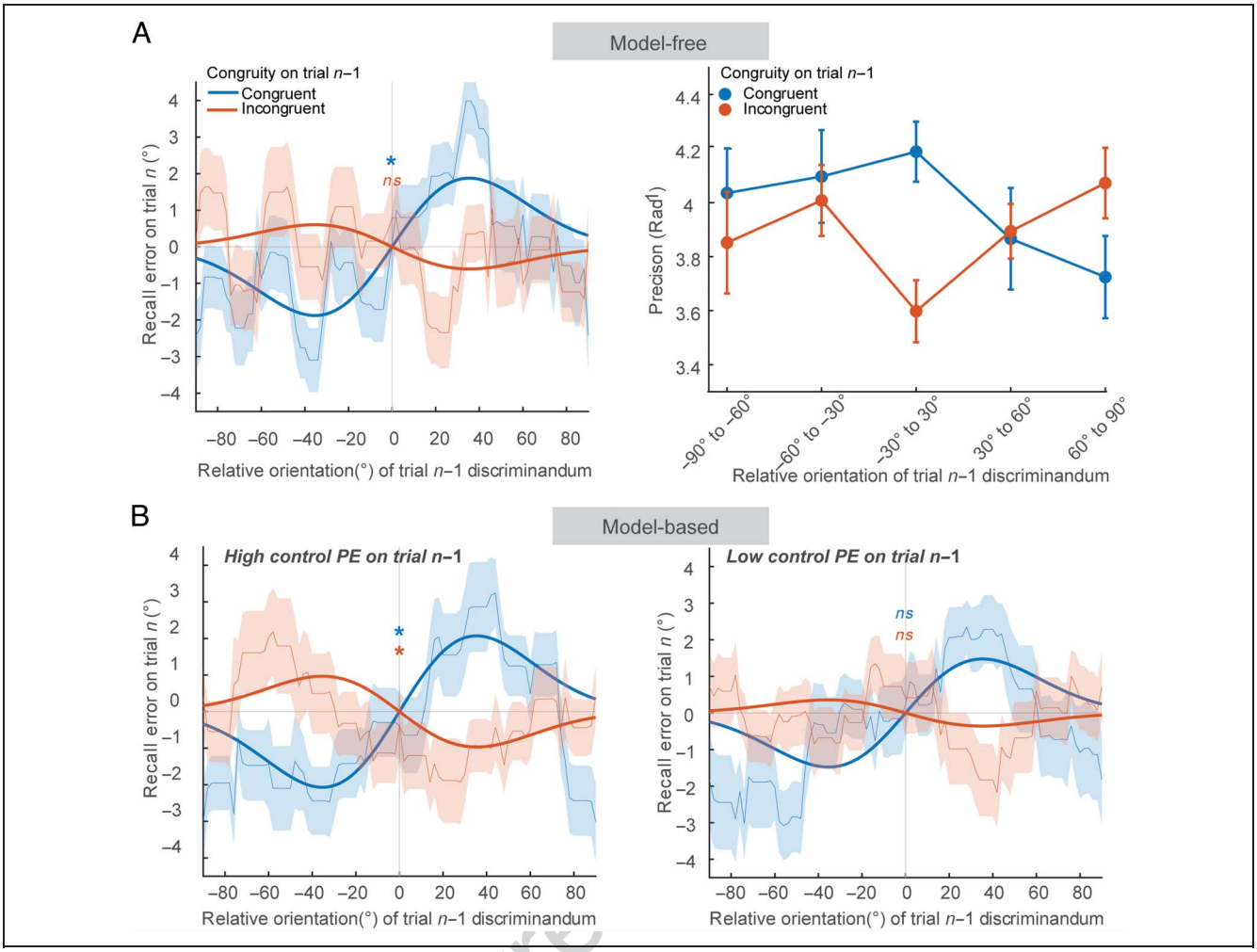


Figure 3. Behavioral effects of conflict adaptation on WM recall. (A) Left-hand column: Serial dependence analysis of the influence of the trial *n*–1 discriminandum on WM recall on trial *n*. After trial *n* data were sorted as trial *n*–1-congruent and trial *n*–1-incongruent, WM recall error on trial *n* was plotted based on the relative orientation of the trial *n* discriminandum. Only the attractive serial bias exerted by congruent discriminanda was statistically reliable. Right-hand column: WM recall precision on trial *n* plotted as a function of the relative angular distance from the discriminandum on trial *n*–1, revealing that an orientation-dependent serial bias of the discriminandum was largely opposite as a function of congruity. (B) Influence of the trial *n*–1 discriminandum on WM recall on trial *n* when trials are sorted by the *control PE* parameter from fits to the FCM. Left-hand column: for the tertiles of congruent and incongruent discriminanda with the highest *control PE*s. Congruent discriminanda exerted an attractive bias on WM recall on trial *n*, and incongruent discriminanda exerted a repulsive bias on WM recall on trial *n*. Right-hand column: For the tertiles of congruent and incongruent discriminanda with the lowest *control PE, neither* congruent nor incongruent discriminanda exerted a significant serial bias. * Indicates two-sided $p < .05$ against 0.

conflict (and, therefore, preparatory control) is low, the discriminandum is (unexpectedly) incongruent with the memorandum. This intuition is better captured by the FCM than by the simple sorting of trials by congruity (see the Model-free Analyses, WM Recall section, above) because the *control PE* parameter factors is "the element of surprise": It is instances of discrepancy between expected conflict and observed conflict that generate large control *PE*s. Indeed, a prediction arising from Teng and colleagues (2022) and tested here is that the *control PE* recruits reflexive control, in a manner such that the magnitude of the *control PE* predicts the amount of reflexive control that is recruited.

For these analyses, trials were sorted by congruity and by the amount of the *control PE* parameter. Whereas congruent discriminanda eliciting high *control PE*s on trial *n*–1 exerted the expected attractive bias on WM recall on trial *n* ($a = 2.07°$, $p < .001$, 95% CI [0.96°, 3.25°], bootstrapped), incongruent discriminanda eliciting high *control PE*s on trial *n*–1 exerted a repulsive bias on WM recall on trial *n* ($a = -1.17°$, $p = .043$, 95% CI [−2.42°, −0.08°], bootstrapped), and the two curves differed significantly from each other ($p = .002$, bootstrapped; Figure 3B, left column). For the tertile of trials that were preceded by a trial with a congruent discriminandum that elicited a low *control PE*, the attractive serial bias exerted by the congruent discriminandum only trended toward significance ($a = 1.47°$, $p = .058$, 95% CI [0.16°, 2.67°], bootstrapped), and for the tertile of trials that were preceded by a trial with an incongruent

Downloaded from http://direct.mit.edu/jocn/article-pdf/doi/10.1162/jocn_a_02310/2501480/jocn_a_02310.pdf?casa_token=VLDfFMCJQ6UAAAAA:63PtRA1U-xfVX1sUK_IgFiFVeddBeHgHOMm8ObNCYPzfPDY8tiN688x0-HCEDB4viY_Go by University of Wisconsin, Madison, Jacqueline Fulvio on 23 April

discriminandum that elicited a low *control PE*, no serial bias was evident ($a = -0.36°$, $p = .57$, 95% CI [−1.28°, 0.68°], bootstrapped; Figure 3B, right column).

Results of the analysis of WM recall precision on trial n as a function of the relative angular distance from the discriminandum on trial n–1 (Figure 3A; right column) are consistent with the possibility that the reflexive control triggered by an incongruent discriminandum results in feature-specific suppression of the representation of that discriminandum. Previous work has documented impairments in performance when a suppressed feature reappears as a visual search target (e.g., Gaspelin, Luck, & Kelley, 2015; Sawaki & Luck, 2010) or a WM sample (Teng et al., 2022). Therefore, we examined whether the precision of WM recall was modulated by the congruity of the discriminandum from the preceding trial in an orientation-dependent manner. To do this, we first sorted trials by congruity on trial n–1 and then further divided them into five bins according to the relative orientation difference between the discriminandum from trial n–1 and the memorandum from trial n: −90° to −60°, −60° to −30°, 0° to 30°, 30° to 60°, and 60° to 90° (Figure 3A, right). For statistical analysis, negative orientation distances were folded over to the positive side to increase power, resulting in three bins: 0° to ±30°, ±(30° to 60°), and ±(60° to 90°). A repeated-measures ANOVA with factors of trial n congruity and orientation distance revealed a significant interaction between the two factors, $F(2, 48) = 6.04$, $p = .005$. Post hoc comparisons indicated that when a congruent discriminandum was followed by a memorandum with a similar orientation (within the 0° to ±30° range) on the next trial n, the recall precision of that memorandum was enhanced ($p = .04$ compared with ±[30° to 60°]; $p = .019$ compared with ±[60° to 90°], Bonferroni corrected). For an incongruent discriminandum, in contrast, there was a trend suggesting that recall precision was worse when the memorandum from trial n was similar in orientation to the discriminandum from trial n–1, although none of the pairwise comparisons survived Bonferroni correction. Note that we did not perform this orientation-dependent analysis in relation to the FCM parameter estimates due to the limited number of trials per bin (less than 20), which would yield unreliable estimates of recall precision.

EEG Results

Preparatory Control Elevates Pretrial Theta-band Power at Frontal Midline Electrodes and Enhances WM Representation

Within the framework of the FCM, the level of preparatory control (f) during performance of the WM + discrimination dual task is updated on each trial by the *control PE* elicited by the discriminandum on trial n–1, and this new level of f persists until it is updated by the *control PE* on trial n. Thus, it should manifest as a signal that persists from shortly after the processing of the discriminandum on trial n–1 until it is updated by the *control PE* triggered by the discriminandum on trial n. To search for a neural correlate of such a signal, we focused on theta-band oscillations within a time window of interest defined as −800 to 1000 msec relative to trial onset. The data were baseline corrected from −1000 to −800 msec relative to trial onset, and theta-band power was extracted as the 4- to 7-Hz band-passed signal from all electrodes. The time series signal was regressed against f, and the resultant regressor weights run through a cluster-based permutation test in 2-D (channel × time point). This analysis revealed a cluster of electrodes along the frontal midline that exhibited significant regression weights for f. The significant loading on frontal midline theta started to emerge around −460 msec and lasted until −40 msec relative to the onset of the trial (Figure 4A–B, $p = .03$). Figure 4C illustrates the spectral difference between the tertile of trials with the highest f and the tertile of trials with the lowest f at electrode Cz. (In addition, regression analyses on alpha-band (8–13 Hz) and beta-band (14–30 Hz) activities during ITI failed to identify any significant clusters with significant correlation of predicted conflict f.)

To investigate how the level of preparatory control may influence WM processing, we conducted a multivariate decoding analysis of the representation of the orientation of the memorandum. On trials with high f, reliable decoding of the memorandum emerged roughly 100 msec after stimulus onset ($p = .038$, cluster corrected). On trials with low f, however, above-baseline decoding did not emerge until roughly 400 msec following stimulus onset ($p = .050$, cluster corrected). Furthermore, decoding accuracy was significantly higher on high f trials compared with low f trials during the early stages of memorandum presentation (80–100 msec after memorandum onset, $p = .039$, cluster corrected; Figure 5B and C).

The Interplay between Preparatory and Reflexive Control during Perceptual Discrimination

Within the framework of the FCM, discriminandum onset generates the binary value of observed trial congruity (o), and the difference between o and f determines the sign and amplitude of the resultant *control PE*, which, in turn, triggers the engagement of the reflexive control that may be needed to control the consequences of memorandum-discriminandum congruity. Because the *control PE* is a transient signal, we focused on discriminandum-locked ERPs to examine these dynamics (−200 to 1000 msec relative to discriminandum onset; baseline corrected to 200 msec before discriminandum onset). Regression analyses were performed with f, *control PE*, and trial congruity as regressors against EEG voltage (i.e., the ERP), revealing distinct timing and scalp topographies for each regressor (Figure 6). Beginning with f, its influence emerged early (152 to 256 msec after discriminandum onset, $p = .03$, cluster corrected), with a significant cluster forming around the central electrodes that showed increased

Uncorrected Proof

Downloaded from http://direct.mit.edu/jocn/article-pdf/doi/10.1162/jocn_a_02310/2501480/jocn_a_02310.pdf?casa_token=VLDfMCJQ6UAAAAA:63PtRA1U-xfVX1sUK_IgFiFVeddBeHgHOMm8ObNCYPzfPDY8tiN688x0-HCEDB4viY_Go by University of Wisconsin, Madison, Jacqueline Fulvio on 23 April

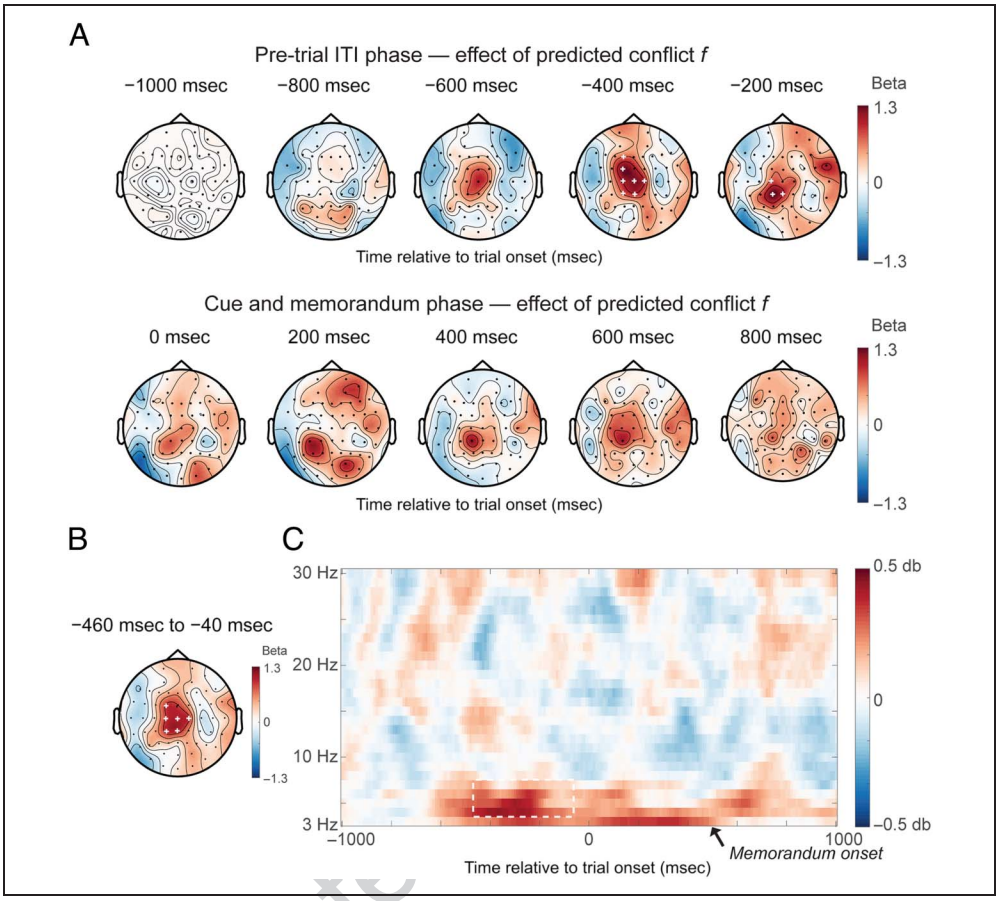


Figure 4. Neural signature of preparatory control – pretrial theta power at frontal midline electrodes scales with the level of expected conflict. (A) Scalp topography of regression weights for the *f* parameter from the FCM derived from the regression analysis on theta-band (4–7 Hz) activity. Top row: 200-msec windows of the final second of ITI before onset of trial *n* (with *f* having been updated by the discriminandum-triggered *control PE* from trial *n*–1); bottom row: 200-msec windows of the first second of trial *n*. (B) Topography of the cluster of electrodes with significant correlation of *f* with theta-band power between −460 msec and −40 msec relative to trial onset. (C) Time–frequency plot at electrodes from the identified cluster (Cz, C1, C2, FC1, CP1, and CPz) showing an increase of theta-band power before the onset of the trial. The white dashed outline indicates the significant time points from the cluster in B identified through the cluster-based permutation test.

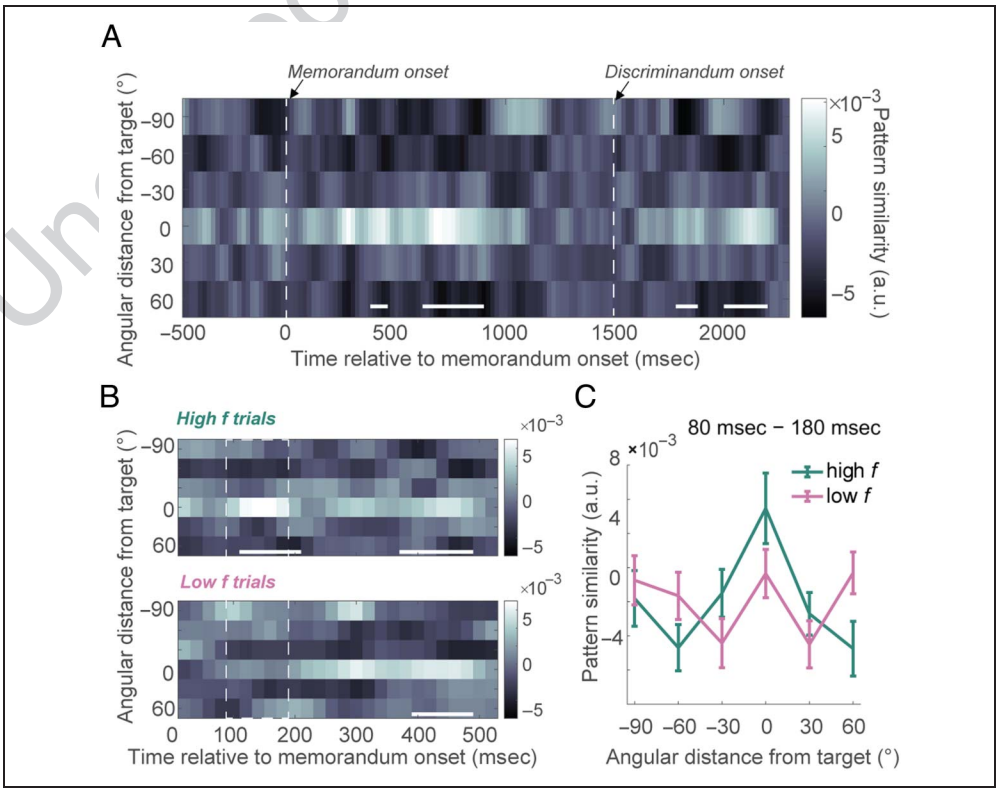


Figure 5. Influence of preparatory control on the neural representation of the memorandum. (A) Time course of memorandum decoding averaged across all trials. Similarity patterns (Mahalanobis distance, sign-reversed) from a single test trial to the training template were mean-centered and averaged for each angular distance and time point. Decoding was successful after the onset of the memorandum and the onset of the discriminandum. Solid white lines mark statistically significant time periods (two-tailed $p < .05$, cluster corrected). (B) Average similarity patterns for memorandum orientation on high *f* and low *f* trials. Whereas on high *f* trials the memorandum was quickly decodable after its onset, on low *f* trials, it was not decodable until around 400 msec. Solid white lines indicate statistically significant time periods (against zero; $p < .05$, cluster corrected). The dashed white lines mark time points at which high and low f trials are statistically different. (C) Similarity patterns averaged over 80–180 msec identified in B illustrate more robust stimulus representation on high *f* trials than low *f* trials.

Downloaded from http://direct.mit.edu/jocn/article-pdf/doi/10.1162/jocn_a_02310/2501480/jocn_a_02310.pdf?casa_token=VLDfFMCJQ6UAAAAA:63PiRA1U-xfVX1sUK_IgFiFVeddBeHgHOMm8ObNCYPzfPDY8tiN688x0-HCEDB4wiY_Go by University of Wisconsin, Madison, Jacqueline Fulvio on 23 April

positivity with higher values of *f* (Figure 6A). The influence of *o* was seen at posterior electrodes, with incongruent trials displaying increased negativity from 444 to 800 msec ($p = .02$, cluster corrected; Figure 6B). Finally the *control PE* elicited a left frontal positivity from 420 to 742 msec post discriminandum onset ($p = .04$, cluster corrected; Figure 6C), thereby appearing at a similar time as the effect of *o*. The right-hand column in Figure 6 illustrates the difference between the tertile of trials with the highest and lowest *f*, congruent and incongruent trials, and the tertile of trials with the highest and lowest *control PE*. These results were consistent with the hypothesis that preparatory control is engaged during the early processing of the discriminandum, and the *control PE* emerges upon the processing of discriminandum congruity. The same regression analysis was performed against spectral power and identified clusters of left anterior and of right posterior electrodes whose activity was modulated by congruity. These clusters showed increased theta-band power associated with incongruent stimuli 780–960 msec following the discriminandum onset ($p = .01$, cluster-corrected; Figure A3). None of the other effects were significant.

Neural Consequences of Reflexive Control

The behavioral data from this study and from Teng and colleagues (2022) indicate that an incongruent discriminandum impairs recall of the memorandum at the end of the trial. We have previously hypothesized that the reflexive control triggered by an unexpectedly incongruent discriminandum has the effect of suppressing its neural representation, so as to minimize its potential to disrupt WM recall (Teng et al., 2022). One of the behavioral hallmarks of such suppression may be the repulsive serial bias exerted by incongruent discriminanda that generate high *control PE*s (Figure 3; Teng et al., 2022), an effect that has also been observed for items that are actively removed from WM (Shan & Postle, 2022). Here, we assessed neural evidence for suppression-by-reflexive control by examining the neural coding of discriminanda in different conditions.

Proof

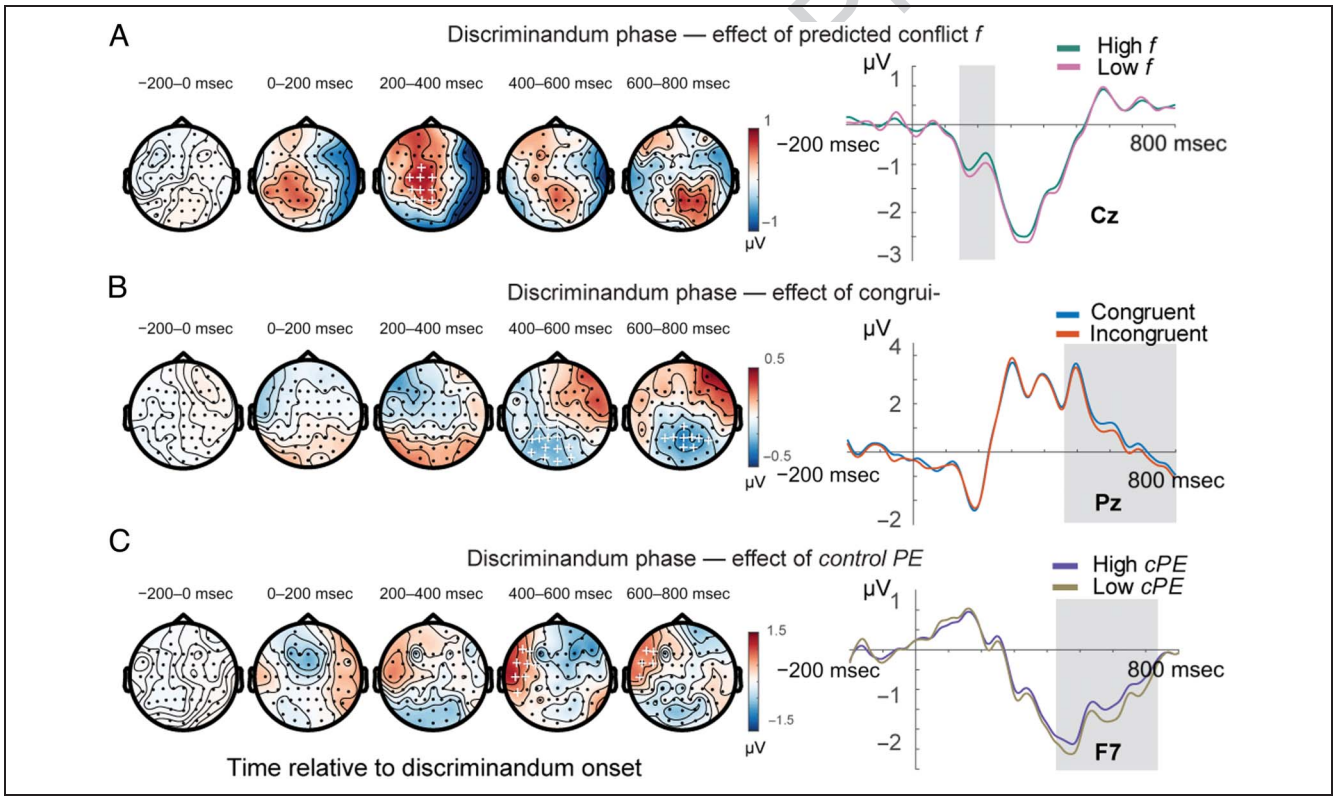


Figure 6. The interplay between preparatory and reflexive control during the discrimination task. (A) Left: scalp topography of regression weight for predicted conflict (*f*). Discriminandum-locked ERP from a large cluster of central electrodes was sensitive to *f* from ~150–250 msec. Right: ERPs plotted from Cz, composed of trials with the highest and lowest tertile of values of *f*. (B) Left: scalp topography of regression weight for congruity between the memorandum and the discriminandum (*o*), revealing a cluster of posterior electrodes that was modulated by congruity beginning at ~450 msec after discriminandum onset. Right: ERPs plotted from Pz for congruent and incongruent trials. (C) Left: scalp topography of regression weights for *control PE*, revealing a cluster of left-frontal electrodes that were significantly modulated by *control PE* from ~420–742 msec following discriminandum onset. Right: ERPs plotted from F7 for the highest and lowest tertile of trials for *control PE*. Gray shaded areas indicate time points of significant difference identified by the regression analysis (two-tailed $p < .05$, cluster-corrected). The timing of these effects is consistent with the temporal dynamics hypothesized by the FCM: an early engagement of preparatory control (*f*; because it was present at this level before discriminandum onset), and subsequent comparison of observed conflict (*o*) against *f* to generate the *control PE*. (See Figures 6–1 for the regression analysis performed on averaged band power of theta, alpha, and beta around the discriminandum onset.)

Downloaded from http://direct.mit.edu/jocn/article-pdf/doi/10.1162/jocn_a_02310/2501480/jocn_a_02310.pdf?casa_token=VLDfMCJQ6UAAAAA:63PiRA1U-xfVX1sUK_IgFiFVeddBeHgHOMm8ObNCYPzfPDY8tiN688x0-HCEDB4wiY_Go by University of Wisconsin, Madison, Jacqueline Fulvio on 23 April

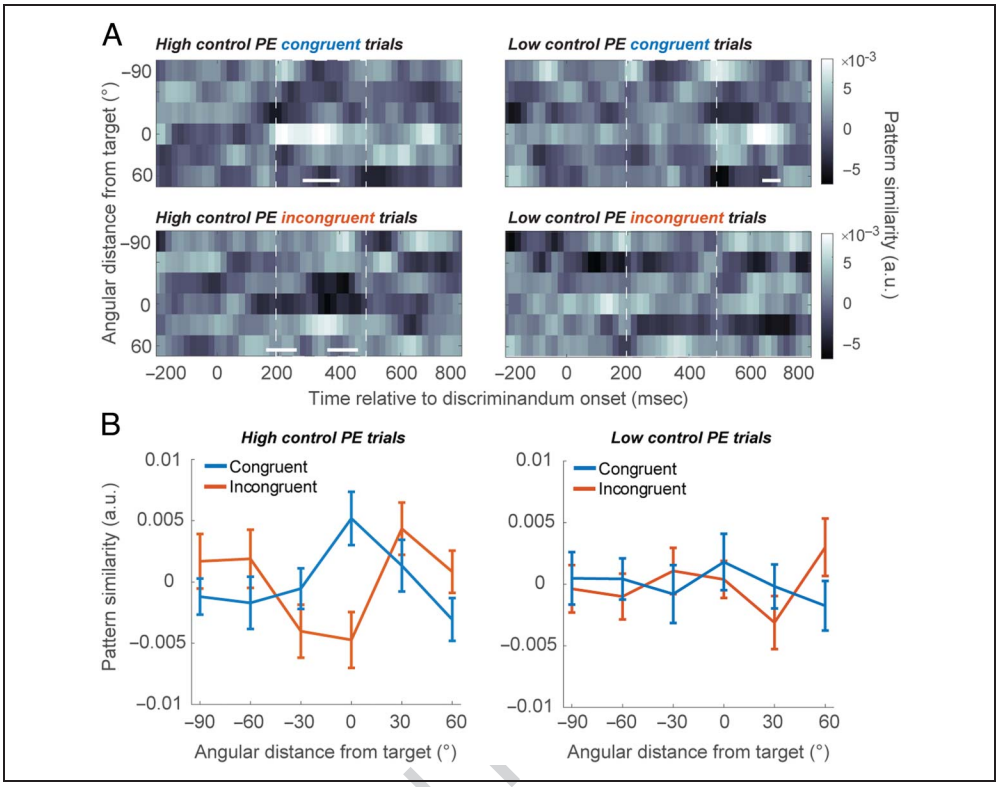


Figure 7. Neural evidence of reflexive suppression as a consequence of control predictor error. (A) Time-resolved decoding results for the discriminandum. For high *control PE* trials, congruent discriminanda were represented in the same active format as the memoranda, whereas incongruent discriminanda were represented in a reversed format, consistent with suppression. Solid white lines indicate significant decoding (two-sided $p < .005$, cluster-corrected); dashed white lines demarcate the span of time when the two differ significantly (i.e., 200–480 msec relative to discriminandum onset). (B) Averaged similarity patterns between 200 and 480 msec for the four trial types.

Figure 7A showed the similarity between the neural representations of the discriminandum and the memorandum. A congruent discriminandum with a high *control PE* could be reliably decoded from the model trained on memorandum orientations (280 to 480 msec relative to discriminandum onset, two-sided $p = .035$, cluster-corrected). Critically, the incongruent discriminandum with a high *control PE* was associated with a "flipped" representation where the similarity pattern had the lowest value at the target orientation (360 to 460 msec relative to discriminandum onset, two-sided $p = .021$, cluster-corrected), suggesting that its neural representation was flipped relative to its representation when the same orientation served as a memorandum. The congruent and incongruent discriminanda significantly differed from each other during the discriminandum-locked time window of 200 to 480 msec. For low *control PE* trials, a brief period of decodability of the congruent discriminandum was observed at a later time point with a positive tuning profile (640 to 700 msec, two-sided $p = .048$, cluster-corrected). A further statistical comparison between high and low *control PE* incongruent discriminanda revealed a significant difference during the time window of 200 to 480 msec (two-sided $p = .043$, bootstrapped). Thus, when testing with the model trained on memorandum orientations, incongruent discriminanda with the same orientation values were reconstructed differently depending on whether their incongruity was expected or not. Figure 7B depicts the tuning profiles for those four trial types during the time window when congruent and incongruent discriminanda on high *control PE* trials differed significantly (200 to 480 msec). These findings align with the behavioral serial dependence results and provide direct neural evidence for the active involvement of reflexive control, whereby incongruent perceptual input is actively suppressed to resolve conflicts between it and the contents of WM. In addition, we confirmed that orientation values did not vary significantly across tertiles of *f* and absolute *control PE*, indicating that these results cannot be explained by systematic covariation between orientation values and model parameters.

DISCUSSION

In this study, we investigated the role of two modes of cognitive control in mitigating interference between perception and WM. Although previous research has demonstrated the engagement of control during WM–perception interactions, here, we used model-based EEG to track the temporal dynamics of these hypothetically distinct mechanisms, as well as their influence on perceptual and mnemonic representations.

Our study successfully applied the FCM, originally developed for conflict adaptation tasks that lack an explicit demand on WM (e.g., the Stroop task), to tease apart preparatory from reflexive control on a trial-by-trial basis. After establishing that our behavioral results replicated the findings from previous work (Teng et al., 2022)—that parameter estimates from the FCM explain key effects of the control of conflict on performance on the dual WM-plus-discrimination task—we directly linked these parameter estimates to EEG signals via multiple-regression analysis. First, we

Uncorrected

Downloaded from http://direct.mit.edu/jocn/article-pdf/doi/10.1162/jocn_a_02310/2501480/jocn_a_02310.pdf?casa_token=VLDfMCJQ6UAAAAA:63PiRA1U-xfVX1sUK_IgFiFVeddBeHgHOMm8ObNCYPzfPDY8tiN688x0-HCEDB4viY_Go by University of Wisconsin, Madison, Jacqueline Fulvio on 23 April

observed that theta-band power at frontal midline electrodes closely tracked the level of preparatory control, a signal that, updated by recent trial history, was most prominent before trial onset. With the onset of the trial, the effects of preparatory control were evident in the neural reconstruction (i.e., the representational strength) of the memorandum and in early components of the discriminandum-locked ERP. Furthermore, EEG components tracking the effect of congruity and the *control PE* emerged around the same time, suggesting dynamic interactions that give rise to reflexive control. Although context-dependent shifts between proactive and reactive control have been broadly established, studies have typically manipulated conflict either at the list-wise or at the item-specific level to create task contexts that favor one mode of control over the other (e.g., Braver et al., 2021; Gonthier, Braver, & Bugg, 2016; Gonthier, Macnamara, Chow, Conway, & Braver, 2016; Bugg & Hutchison, 2013; Braver, Paxton, Locke, & Barch, 2009). To our knowledge, there has been only limited behavioral evidence to date that examines the simultaneous engagement of the two modes of control (Mäki-Marttunen, Hagen, & Espeseth, 2019). The results presented here provide a novel demonstration of the concurrent recruitment of preparatory and reflexive control at the single-trial level and illustrate how these two modes of control interact, engaging distinct neural processes to minimize conflict-induced interference.

In addition, we note that there is ongoing debate in the literature about whether and how people generate expectations about upcoming conflict in tasks with random trial sequences. For example, some studies found evidence suggesting people do not explicitly expect congruency to repeat across consecutive trials (Duthoo et al., 2014). However, the presence of reliable conflict adaptation effects even in tasks with 50% proportion congruency (e.g., Weissman et al., 2014) indicates that control adjustments occur regardless of explicit expectations. Indeed, determining that a sequence is truly random would require the system to first attempt to learn its temporal statistics—the very process that would detect patterns if they existed. Here, we provided further evidence that the preparatory form of control can be engaged in the absence of a predictable pattern or structure in congruity, highlighting the participants' flexible adjustment of control based on recent experiences.

The finding linking theta-band power at frontal midline electrodes with the level of preparatory control is consistent with the broader literature on cognitive control, and one influential view is that "frontal-midline theta" (FMθ) represents the general need for and anticipation of the need for control (e.g., Kaiser, Iliopoulos, Steinmassl, & Schütz-Bosbach, 2022; Messel et al., 2021; Cooper et al., 2019; Kaiser, Simon, Sauseng, & Schütz-Bosbach, 2019; Chang, Chen, & Li, 2017; Cavanagh & Frank, 2014; Cavanagh, Zambrano-Vazquez, & Allen, 2012). For example, one recent study showed that in a conflict action paradigm, presenting a predictive cue to indicate upcoming conflict before the action signal led to increase in FMθ power during the preparation phase, indicating a role in conflict preparation (Messel et al., 2021). Here, our results showed that trial-by-trial levels of FMθ power fluctuated with predicted need for the control of conflict in an anticipatory fashion, before trial onset, thus providing additional support for the link between FMθ and anticipatory control. Furthermore, FMθ has been studied primarily in task contexts involving response conflict or response inhibition (e.g., Messel et al., 2021; Chang et al., 2017). In the present study, incongruity occurred at the level of stimulus representation, putting the loci of conflict not on the response but rather on the processes of perceptual decision making and WM recall. Our findings thus suggest that FMθ signaling may serve as a general mechanism for regulating top–down bias on sensory and motor processing networks, even in the absence of overt response conflict. Interestingly, one of the EEG results presented here has also provided evidence for a neural effect of preparatory control for which we have not found a behavioral correlate. This relates to the fact that although WM recall accuracy suffers on incongruent trials, this effect does not seem to be mediated by control (model-free: it is insensitive to congruity on trial n–1; model-based: it is insensitive to f). Nonetheless, decoding of the memorandum is stronger at high levels of f, an effect that would typically be interpreted as evidence for stronger encoding into WM. Although this requires further study, it is possible that any control-related enhancement of encoding into WM in this WM-plus-discrimination task is diluted by the events that intervene between encoding and recall.

Turning to the *control PE*, our analyses identified a cluster of left anterior electrodes for which voltage scaled positively with the magnitude of the *control PE*. This effect is consistent with past research that has implicated lateral pFC in conflict resolution (e.g., Burgess & Braver, 2010; Feredoes et al., 2006) and in learning-dependent control (Jiang et al., 2015). In a color-naming Stroop task that compared the coding schemes in DLPFC and dorsomedial frontal cortex (DMFC), DMFC was found to encode the trial incongruity (control demand) and DLPFC both target incongruity-related information. Importantly, stronger target coding in DLPFC was associated with a reduced Stroop interference (Freund, Bugg, & Braver, 2021), consistent with the proposed functionally dissociated roles of DMFC in control demand monitoring and of DLPFC in control implementation (Shenhav, Botvinick, & Cohen, 2013; Braver & Cohen, 2000). For WM, conflict processing has been studied extensively with a "recent-negative probes" procedure that manipulates proactive interference from previously encountered information (e.g., Jonides & Nee, 2006; Jonides, Smith, Marshuetz, Koeppe, & Reuter-Lorenz, 1998; Monsell, 1978). The common finding with this procedure is that a probe that was not been a member of the current trial's memory set but was a member of the previous trial's memory set takes longer to reject

Uncorrected Proof

Downloaded from http://direct.mit.edu/jocn/article-pdf/doi/10.1162/jocn_a_02310/2501480/jocn_a_02310.pdf?casa_token=VLDfMCJQ6UAAAAA:63PtRA1U-xfVX1sUK_IgFiFVeddBeHgHOMm8ObNCYPzfPDY8tiN688x0-HCEDB4viY_Go by University of Wisconsin, Madison, Jacqueline Fulvio on 23 April

and elicits increased activity in left inferior frontal gyrus (IFG; e.g., Postle et al., 2004; Jonides et al., 1998). Furthermore, repetitive TMS applied to left IFG concurrent with probe onset selectively exaggerates the RT cost associated with recent-negative probes (Feredoes & Postle, 2009; Feredoes et al., 2006).

In the context of the FCM, model-based analyses of neuroimaging data have shown that, in a modified Stroop task, activity in left anterior insula and adjacent IFG is linked both to the magnitude of the *control PE* and to the flexible learning rate in the FCM that governs the amplitude of the change to f that is produced by a given value of the *control PE* (i.e., the α parameter from the FCM, which we did not manipulate in the present study; Jiang et al., 2015). Furthermore, repetitive TMS of left lateral pFC eliminates behavioral effects of phasic and tonic adaptation to changes in the probability of incongruent trials, an effect interpreted as a disruption of the processes by which the *control PE* drives updates to f (phasic adaptation) and to α (tonic adaptation; Muhle-Karbe et al., 2018). Thus, the present results are consistent with the idea that the left lateral pFC contributes to the processing of the *control PE* that leads to flexible adjustments in the control of preparatory control.

The present results also provide novel evidence about mechanisms of reflexive control. Recall that the need for reflexive control of incongruent discriminanda is illustrated by the fact that incongruity between memoranda and discriminanda is associated with lower WM recall precision. Behaviorally, the fact that discriminanda that generate large *control PE*s also exert stronger serial biases (Figure 3B) has been interpreted as evidence for a role for the *control PE* in triggering reflexive control, so as to minimize the interference from incongruent discriminanda on recall of the memoranda (Teng et al., 2022). Here, analyses of the concurrently recorded EEG data provide neural evidence for this link: Multivariate decoding of discriminanda from the EEG data reveals that, on high *control PE* trials, the neural representations of incongruent discriminanda are suppressed below baseline, a presumed immediate response to conflict. This finding confirms that the control *PE* is the primary driver of reflexive control and its associated behavioral effects on the subsequent trial. It further suggests a neural explanation for why the serial bias exerted by an item from trial n–1 on WM recall on trial n will be either repulsive or attractive as a function of whether its representation needs to be suppressed (Shan & Postle, 2022; Teng et al., 2022). It also supports the interesting possibility that active suppression may be a common mechanism that the visual system employs to reduce interference from task-irrelevant information in many contexts, including visual perception (e.g., Gaspelin et al., 2015) and visual WM (Shan & Postle, 2022; Teng et al., 2022). Although model-based analyses implicate the *control PE* as a signal that can trigger reflexive control, understanding the source(s) of the reflexive control signal itself will require further work.

What is the nature of "preparatory control" and "reflexive control" as operationalized by the f and control PE parameters of the FCM? The fact that f predicts pretrial frontal midline theta suggests similarities between "preparatory control" and the construct of anticipatory control from, for example, Kaiser and colleagues (2019, 2022), Messel and colleagues (2021), Cooper and colleagues (2019), Chang and colleagues (2017), Cavanagh and Frank (2014), and Cavanagh and colleagues (2012). In the context of our dual task, the influence of preparatory control on behavior was to implement conflict adaptation, and its influence on EEG correlates was the strengthen the encoding of the memorandum. The influence of reflexive control on our dual task, as mentioned above, was seen in the processing of the discriminandum: Its immediate effect was to suppress the EEG representation of the discriminandum, and its downstream effects on behavior were seen in the orientation-specific disruption of recall on the current trial and in the repulsive serial bias exerted on recall on the subsequent trial. It is noteworthy that the suppression effect in the EEG was not simply to weaken the strength of the discriminandum but to actually transform it into an inverted, or "negative," version of itself. Although this effect was only observable in the EEG while the discriminandum was on the screen, a persistence of this inversion could account for the orientation specificity of the behavioral effects on recall.

To conclude, our results demonstrate the trial-by-trial neural dynamics associated with the interplay between the adaptive adjustment of preparatory control and the recruitment of reflexive control in the control interference between WM and perception. Future work is needed to clarify points of commonality and of difference between the well-established constructs of proactive and reactive control as understood in the dual-processes framework (e.g., Braver et al., 2009) and the preparatory and reflexive processes studied here. In addition to the control signals themselves, our results also characterized the consequences of these control processes for the neural representation of stimulus information in the visual system. Understanding the neural dynamics of distinct control mechanisms for conflict resolution is essential for gaining insights into flexible and adaptive human behavior.

Uncorrected Proof

Downloaded from http://direct.mit.edu/jocn/article-pdf/doi/10.1162/jocn_a_02310/2501480/jocn_a_02310.pdf?casa_token=VLDfMCJQ6UAAAAA:63PtRA1U-xfVX1sUK_IgFiFVeddBeHgHOMm8ObNCYPzfPDY8tiN688x0-HCEDB4viY_Go by University of Wisconsin, Madison, Jacqueline Fulvio on 23 April

APPENDICES

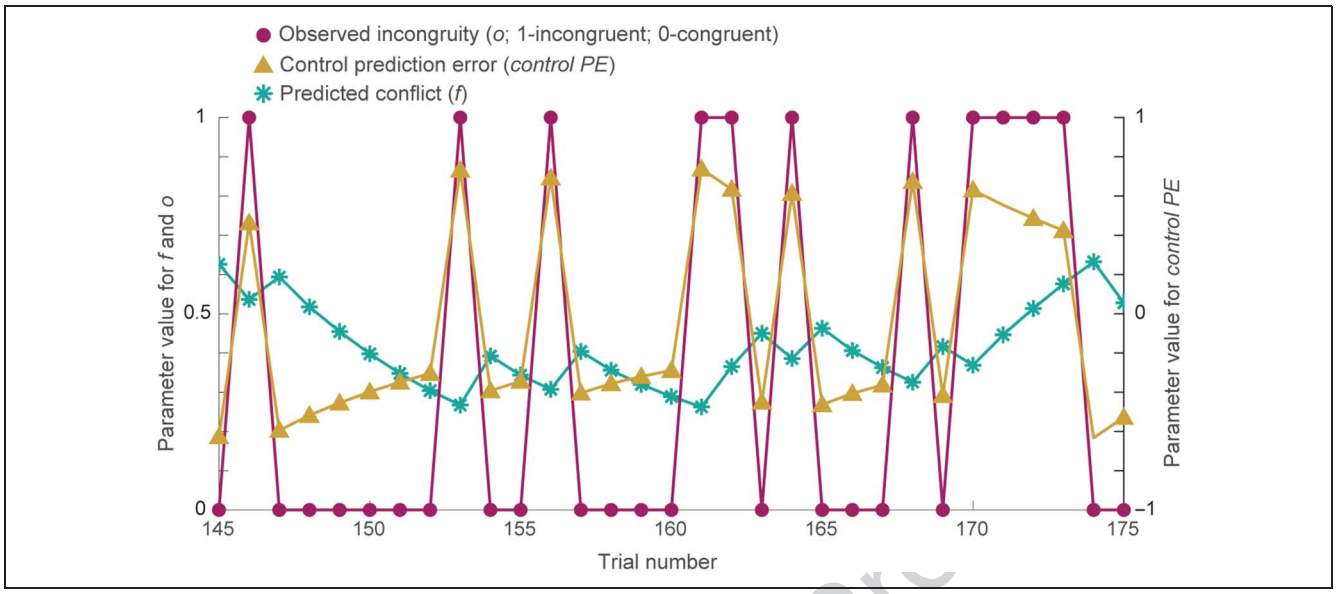


Figure A1. Parameter estimates for 30 consecutive trials from an example participant with the FCM. Purple circles: observed incongruity (o), with 1 representing incongruent and 0 representing congruent. Green asterisks: Predicted conflict f, larger values represent higher predicted conflict (continuous value between 0 and 1). Yellow triangles: signed control prediction error (*control PE*; continuous value between −1 and 1), calculated as the difference between the observed congruity and the predicted conflict ($o - f$). Note that with a series of consecutive congruent trials (e.g., Trials 147–152), there is a steady decrease of the parameter estimate of predicted conflict f. In addition, on those same consecutive trials with the same congruity, because each one is "less surprising," the value of the *control PE* gradually converges toward 0.

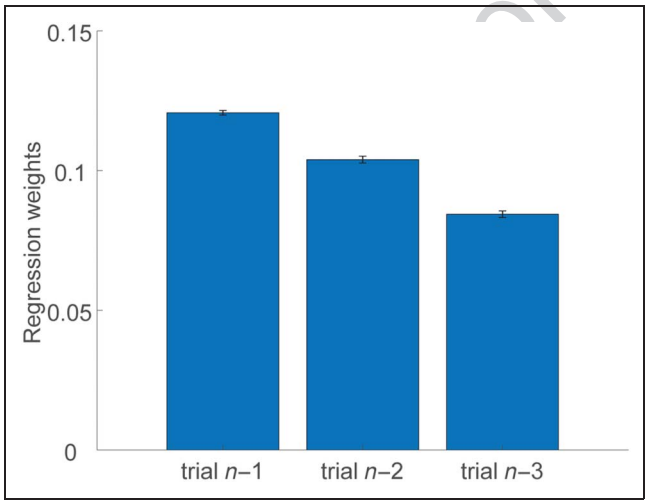


Figure A2. To elucidate the influence of previous trials' congruity on the parameter estimate of predicted conflict f, we performed a regression analysis on f with regressors of trial congruity on trials n–3, n–2, and n–1. In the figure, regression weights were plotted for the three regressors. This demonstrates that f is influenced not only by the most recent trial but also by trials that are farther out. Thus, f captures a broader swath of task history, consistent with a preparatory control mechanism that integrates recent experiences to optimize performance.

Downloaded from http://direct.mit.edu/jocn/article-pdf/doi/10.1162/jocn_a_02310/2501480/jocn_a_02310.pdf?casa_token=VLDfMCJQ6UAAAAA:63PiRA1U-xfVX1sUK_IgFiFVeddBeHgHCMm8ObNCYPzfPDY8tiN688x0-HCEDB4wIY_Go by University of Wisconsin, Madison, Jacqueline Fulvio on 23 April

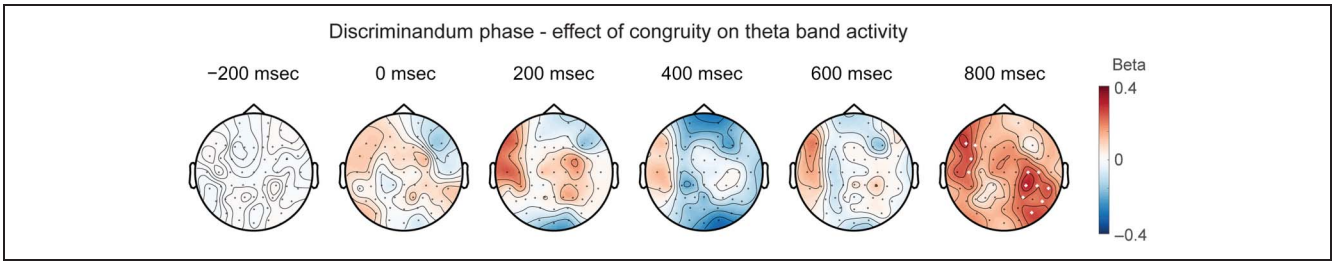


Figure A3. Regression analysis was performed on averaged band power of theta, alpha, and beta around the discriminandum onset, with regressors of *f*, absolute *control PE*, and congruity. Regression weights for trial congruity revealed two significant clusters, a left anterior cluster and a right posterior cluster, that showed increased theta-band power with incongruent discriminanda than congruent ones. None of the other effects were significant.

Corresponding author: Chunyue Teng, Department of Neuroscience, Lawrence University, Appleton, WI 54911, USA, e-mail: chunyue.teng@lawrence.edu.

Data Availability Statement

Data will be made available upon reasonable request.

Author Contributions

Chunyue Teng: Conceptualization; Data curation; Formal analysis; Investigation; Methodology; Software; Visualization; Writing—Original draft; Writing—Review & editing. Jacqueline M. Fulvio: Conceptualization; Data curation; Formal analysis; Investigation; Software; Writing—Review & editing. Mattia Pietrelli: Data curation; Formal analysis; Methodology; Writing—Review & editing. Jiefeng Jiang: Conceptualization; Formal analysis; Methodology; Software; Writing—Review & editing. Bradley R. Postle: Conceptualization; Funding acquisition; Investigation; Methodology; Supervision; Writing—Review & editing.

Funding Information

National Institute of Mental Health (https://dx.doi.org/10.13039/100000025), grant number: MH095984.

Diversity in Citation Practices

Retrospective analysis of the citations in every article published in this journal from 2010 to 2021 reveals a persistent pattern of gender imbalance: Although the proportions of authorship teams (categorized by estimated gender identification of first author/last author) publishing in the *Journal of Cognitive Neuroscience* (*JoCN*) during this period were M(an)/M = .407, W(oman)/M = .32, M/W = .115, and W/W = .159, the comparable proportions for the articles that these authorship teams cited were M/M = .549, W/M = .257, M/W = .109, and W/W = .085 (Postle and Fulvio, *JoCN*, 34:1, pp. 1–3). Consequently, *JoCN* encourages all authors to consider gender balance explicitly when selecting which articles to cite and gives them the opportunity to report their article's gender citation balance. The authors of this article report its proportions of citations by gender category to be: M/M = .627; W/M = .254; M/W = .102; W/W = .017.

REFERENCE

Aron, A. R. (2007). The neural basis of inhibition in cognitive control. *Neuroscientist*, *13*, 214–228. https://doi.org/10.1177/1073858407299288, PubMed: 17519365

Baddeley, A. (1992). Working memory. *Science*, *255*, 556–559. https://doi.org/10.1126/science.1736359, PubMed: 1736359

Bae, G.-Y., & Luck, S. J. (2019). What happens to an individual visual working memory representation when it is interrupted? *British Journal of Psychology*, *110*, 268–287. https://doi.org/10.1111/bjop.12339, PubMed: 30069870

Bliss, D. P., Sun, J. J., & D'Esposito, M. (2017). Serial dependence is absent at the time of perception but increases in visual working memory. *Scientific Reports*, *7*, 14739. https://doi.org/10.1038/s41598-017-15199-7, PubMed: 29116132

Boettcher, S. E. P., Gresch, D., Nobre, A. C., & van Ede, F. (2021). Output planning at the input stage in visual working memory. *Science Advances*, *7*, eabe8212. https://doi.org/10.1126/sciadv.abe8212, PubMed: 33762341

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. https://doi.org/10.1037/0033-295X.108.3.624, PubMed: 11488380

Braem, S., Bugg, J. M., Schmidt, J. R., Crump, M. J. C., Weissman, D. H., Notebaert, W., et al. (2019). Measuring adaptive control in conflict tasks. *Trends in Cognitive Sciences*, *23*, 769–783. https://doi.org/10.1016/j.tics.2019.07.002, PubMed: 31331794

Braver, T. S. (2012). The variable nature of cognitive control: A dual mechanisms framework. *Trends in Cognitive Sciences*, *16*, 106–113. https://doi.org/10.1016/j.tics.2011.12.010, PubMed: 22245618

Braver, T. S., & Cohen, J. D. (2000). On the control of control: The role of dopamine in regulating prefrontal function and working memory. In *Control of cognitive processes: Attention and performance XVIII*.

Braver, T. S., Kizhner, A., Tang, R., Freund, M. C., & Etzel, J. A. (2021). The dual mechanisms of cognitive control project. *Journal of Cognitive Neuroscience*, *33*, 1990–2015. https://doi.org/10.1162/jocn_a_01768, PubMed: 34407191

Braver, T. S., Paxton, J. L., Locke, H. S., & Barch, D. M. (2009). Flexible neural mechanisms of cognitive control within human prefrontal cortex. *Proceedings of the National Academy of Sciences, U.S.A.*, *106*, 7351–7356. https://doi.org/10.1073/pnas.0808187106, PubMed: 19380750

Uncorrected Proof

Downloaded from http://direct.mit.edu/jocn/article-pdf/doi/10.1162/jocn_a_02310/2501480/jocn_a_02310.pdf?casa_token=VLDfMCJQ6UAAAAA:63PiRA1U-xfVX1sUK_IgFiFVeddBeHgHOMm8ObNCYPzfPDY8tiN688x0-HCEDB4viY_Go by University of Wisconsin, Madison, Jacqueline Fulvio on 23 April 2021

Brown, J. W., & Braver, T. S. (2005). Learned predictions of error likelihood in the anterior cingulate cortex. *Science*, *307*, 1118–1121. https://doi.org/10.1126/science.1105783, PubMed: 15718473

Bugg, J. M. (2012). Dissociating levels of cognitive control: The case of Stroop interference. *Current Directions in Psychological Science*, *21*, 302–309. https://doi.org/10.1177/0963721412453586

Bugg, J. M., & Crump, M. J. C. (2012). In support of a distinction between voluntary and stimulus-driven control: A review of the literature on proportion congruent effects. *Frontiers in Psychology*, *3*, 367. https://doi.org/10.3389/fpsyg.2012.00367, PubMed: 23060836

Bugg, J. M., & Hutchison, K. A. (2013). Converging evidence for control of color–word Stroop interference at the item level. *Journal of Experimental Psychology: Human Perception and Performance*, *39*, 433–449. https://doi.org/10.1037/a0029145, PubMed: 22845037

Burgess, G. C., & Braver, T. S. (2010). Neural mechanisms of interference control in working memory: Effects of interference expectancy and fluid intelligence. *PLoS One*, *5*, e12861. https://doi.org/10.1371/journal.pone.0012861, PubMed: 20877464

Carlisle, N. B., & Woodman, G. F. (2011). Automatic and strategic effects in the guidance of attention by working memory representations. *Acta Psychologica*, *137*, 217–225. https://doi.org/10.1016/j.actpsy.2010.06.012, PubMed: 20643386

Cavanagh, J. F., & Frank, M. J. (2014). Frontal theta as a mechanism for cognitive control. *Trends in Cognitive Sciences*, *18*, 414–421. https://doi.org/10.1016/j.tics.2014.04.012, PubMed: 24835663

Cavanagh, J. F., Zambrano-Vazquez, L., & Allen, J. J. B. (2012). Theta lingua franca: A common mid-frontal substrate for action monitoring processes. *Psychophysiology*, *49*, 220–238. https://doi.org/10.1111/j.1469-8986.2011.01293.x, PubMed: 22091878

Chang, A., Chen, C., & Li, C. S. R. (2017). Event-related theta power is associated with retrospective duration estimation and stimulus conflict. *Neuroimage*, *149*, 436–444.

Chiew, K. S., & Braver, T. S. (2017). Context processing and cognitive control: From gating models to dual mechanisms. *The Wiley handbook of cognitive control*, 143–166. https://doi.org/10.1002/9781118920497.ch9

Cohen, M. X. (2016). Midfrontal theta tracks action monitoring over multiple interactive time scales. *Neuroimage*, *141*, 262–272. https://doi.org/10.1016/j.neuroimage.2016.07.054, PubMed: 27475291

Cohen, M. X. (2019). *Analyzing neural time series data: Theory and practice*. MIT Press. https://doi.org/10.7551/mitpress/9609.001.0001

Cooper, P. S., Karayanidis, F., McKewen, M., McLellan-Hall, S., Wong, A. S. W., Skippen, P., et al. (2019). Frontal theta predicts specific cognitive control-induced behavioural changes beyond general reaction time slowing. *Neuroimage*, *189*, 130–140. https://doi.org/10.1016/j.neuroimage.2019.01.022, PubMed: 30639331

Delorme, A., & Makeig, S. (2004). EEGLAB: An open source toolbox for analysis of single-trial EEG dynamics including independent component analysis. *Journal of Neuroscience Methods*, *134*, 9–21. https://doi.org/10.1016/j.jneumeth.2003.10.009, PubMed: 15102499

D'Esposito, M., & Postle, B. R. (2015). The cognitive neuroscience of working memory. *Annual Review of Psychology*, *66*, 115–142. https://doi.org/10.1146/annurev-psych-010814-015031, PubMed: 25251486

Ding, Y., Paffen, C. L. E., Naber, M., & Van der Stigchel, S. (2019). Visual working memory and saliency independently influence the priority for access to visual awareness. *Journal of Vision*, *19*, 9. https://doi.org/10.1167/19.11.9, PubMed: 31532471

Egner, T. (Ed.). (2017). *The Wiley handbook of cognitive control*. John Wiley & Sons. https://doi.org/10.1002/9781118920497

Egner, T., & Hirsch, J. (2005). Cognitive control mechanisms resolve conflict through cortical amplification of task-relevant information. *Nature Neuroscience*, *8*, 1784–1790. https://doi.org/10.1038/nn1594, PubMed: 16286928

Fischer, J., & Whitney, D. (2014). Serial dependence in visual perception. *Nature Neuroscience*, *17*, 738–743. https://doi.org/10.1038/nn.3689, PubMed: 24686785

Freund, M. C., Bugg, J. M., & Braver, T. S. (2021). A representational similarity analysis of cognitive control during color–word Stroop. *Journal of Neuroscience*, *41*, 7388–7402. https://doi.org/10.1523/JNEUROSCI.2956-20.2021, PubMed: 34162756

Fritsche, M., Mostert, P., & de Lange, F. P. (2017). Opposite effects of recent history on perception and decision. *Current Biology*, *27*, 590–595. https://doi.org/10.1016/j.cub.2017.01.006, PubMed: 28162897

Fukuda, K., Pereira, A. E., Saito, J. M., Tang, T. Y., Tsubomi, H., & Bae, G. Y. (2022). Working memory content is distorted by its use in perceptual comparisons. *Psychological Science*, *33*, 816–829. https://doi.org/10.1177/09567976211055375, PubMed: 35452332

Gaspelin, N., Luck, S. J., & Kelley, K. (2015). Directing attention to locations and to feature categories: Shifting and stopping visual processing. *Journal of Experimental Psychology: Human Perception and Performance*, *41*, 1388–1403.

Gayet, S., Paffen, C. L. E., & Van der Stigchel, S. (2013). Information matching the content of visual working memory is prioritized for conscious access. *Psychological Science*, *24*, 2472–2480. https://doi.org/10.1177/0956797613495882, PubMed: 24121415

Gonthier, C., Braver, T. S., & Bugg, J. M. (2016). Dissociating proactive and reactive control in the Stroop task. *Memory & Cognition*, *44*, 778–788. https://doi.org/10.3758/s13421-016-0591-1, PubMed: 26861210

Gonthier, C., Macnamara, B. N., Chow, M., Conway, A. R. A., & Braver, T. S. (2016). Inducing proactive control shifts in the AX-CPT. *Frontiers in Psychology*, *7*, 1822. https://doi.org/10.3389/fpsyg.2016.01822, PubMed: 27920741

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*, 480–506. https://doi.org/10.1037/0096-3445.121.4.480, PubMed: 1431740

Jiang, J., Beck, J., Heller, K., & Egner, T. (2015). An insula-frontostriatal network mediates flexible cognitive control by adaptively predicting changing control demands. *Nature Communications*, *6*, 8165. https://doi.org/10.1038/ncomms9165, PubMed: 26391305

Jiang, J., Heller, K., & Egner, T. (2014). Bayesian modeling of flexible cognitive control. *Neuroscience & Biobehavioral Reviews*, *46*, 30–43. https://doi.org/10.1016/j.neubiorev.2014.06.001, PubMed: 24929218

Jonides, J., & Nee, D. E. (2006). Brain mechanisms of proactive interference in working memory. *Neuroscience*, *139*, 181–193. https://doi.org/10.1016/j.neuroscience.2005.06.042, PubMed: 16337090

Jonides, J., Smith, E. E., Marshuetz, C., Koeppe, R. A., & Reuter-Lorenz, P. A. (1998). Inhibition in verbal working memory revealed by brain activation. *Proceedings of the National Academy of Sciences, U.S.A.*, *95*, 8410–8413. https://doi.org/10.1073/pnas.95.14.8410, PubMed: 9653200

Uncorrected Proof

Downloaded from http://direct.mit.edu/jocn/article-pdf/doi/10.1162/jocn_a_02310/2501480/jocn_a_02310.pdf?casa_token=VLDfMCJQ6UAAAAA:63PiRA1U-xfVX1sUK_IgFiFVeddBeHgHOMm8ObNCYPzfPDY8tiN688x0-HCEDB4viY_Go by University of Wisconsin, Madison, Jacqueline Fulvio on 23 April 2024

Kaiser, J., Iliopoulos, P., Steinmassl, K., & Schütz-Bosbach, S. (2022). Preparing for success: Neural frontal theta and posterior alpha dynamics during action preparation predict flexible resolution of cognitive conflicts. *Journal of Cognitive Neuroscience*, *34*, 1070–1089. https://doi.org/10.1162/jocn_a_01846, PubMed: 35286387

Kaiser, J., Simon, N. A., Sauseng, P., & Schütz-Bosbach, S. (2019). Midfrontal neural dynamics distinguish between general control and inhibition-specific processes in the stopping of motor actions. *Scientific Reports*, *9*, 13054. https://doi.org/10.1038/s41598-019-49476-4, PubMed: 31506505

Kiyonaga, A., & Egner, T. (2014). The working memory Stroop effect: When internal representations clash with external stimuli. *Psychological Science*, *25*, 1619–1629. https://doi.org/10.1177/0956797614536739, PubMed: 24958685

Kiyonaga, A., & Egner, T. (2016). Center-surround inhibition in working memory. *Current Biology*, *26*, 64–68. https://doi.org/10.1016/j.cub.2015.11.013, PubMed: 26711496

Kiyonaga, A., Egner, T., & Soto, D. (2012). Cognitive control over working memory biases of selection. *Psychonomic Bulletin & Review*, *19*, 639–646. https://doi.org/10.3758/s13423-012-0253-7, PubMed: 22528872

Lorenc, E. S., Mallett, R., & Lewis-Peacock, J. A. (2021). Distraction in visual working memory: Resistance is not futile. *Trends in Cognitive Sciences*, *25*, 228–239. https://doi.org/10.1016/j.tics.2020.12.004, PubMed: 33397602

Mäki-Marttunen, V., Hagen, T., & Espeseth, T. (2019). Proactive and reactive modes of cognitive control can operate independently and simultaneously. *Acta Psychologica*, *199*, 102891. https://doi.org/10.1016/j.actpsy.2019.102891, PubMed: 31400651

Maris, E., & Oostenveld, R. (2007). Nonparametric statistical testing of EEG- and MEG-data. *Journal of Neuroscience Methods*, *164*, 177–190. https://doi.org/10.1016/j.jneumeth.2007.03.024, PubMed: 17517438

Messel, M. S., Raud, L., Hoff, P. K., Stubberud, J., & Huster, R. J. (2021). Frontal-midline theta reflects different mechanisms associated with proactive and reactive control of inhibition. *Neuroimage*, *241*, 118400. https://doi.org/10.1016/j.neuroimage.2021.118400, PubMed: 34311382

Muhle-Karbe, P. S., Jiang, J., & Egner, T. (2018). Causal evidence for learning-dependent frontal lobe contributions to cognitive control. *Journal of Neuroscience*, *38*, 962–973. https://doi.org/10.1523/JNEUROSCI.1467-17.2017, PubMed: 29229706

Olivers, C. N. L., Meijer, F., & Theeuwes, J. (2006). Feature-1based memory-driven attentional capture: Visual working memory content affects visual attention. *Journal of Experimental Psychology: Human Perception and Performance*, *32*, 1243–1265. https://doi.org/10.1037/0096-1523.32.5.1243, PubMed: 17002535

Oostenveld, R., Fries, P., Maris, E., & Schoffelen, J.-M. (2011). FieldTrip: Open source software for advanced analysis of MEG, EEG, and invasive electrophysiological data. *Computational Intelligence and Neuroscience*, *2011*, 156869. https://doi.org/10.1155/2011/156869, PubMed: 21253357

Pascucci, D., Mancuso, G., Santandrea, E., Della Libera, C., Plomp, G., & Chelazzi, L. (2019). Laws of concatenated perception: Vision goes for novelty, decisions for perseverance. *PLoS Biology*, *17*, e3000144. https://doi.org/10.1371/journal.pbio.3000144, PubMed: 30835720

Rademaker, R. L., Bloem, I. M., De Weerd, P., & Sack, A. T. (2015). The impact of interference on short-term memory for visual orientation. *Journal of Experimental Psychology: Human Perception and Performance*, *41*, 1650–1665. https://doi.org/10.1037/xhp0000110, PubMed: 26371383

Samaha, J., Switzky, M., & Postle, B. R. (2019). Confidence boosts serial dependence in orientation estimation. *Journal of Vision*, *19*, 25. https://doi.org/10.1167/19.4.25, PubMed: 31009526

Savine, A. C., & Braver, T. S. (2010). Motivated cognitive control: Reward incentives modulate preparatory neural activity during task-switching. *Journal of Neuroscience*, *30*, 10294–10305. https://doi.org/10.1523/JNEUROSCI.2052-10.2010, PubMed: 20685974

Shan, J., & Postle, B. R. (2022). The influence of active removal from working memory on serial dependence. *Journal of Cognition*, *5*, 31. https://doi.org/10.5334/joc.222, PubMed: 36072093

Shenhav, A., Botvinick, M. M., & Cohen, J. D. (2013). The expected value of control: An integrative theory of anterior cingulate cortex function. *Neuron*, *79*, 217–240. https://doi.org/10.1016/j.neuron.2013.07.007, PubMed: 23889930

Soto, D., Heinke, D., Humphreys, G. W., & Blanco, M. J. (2005). Early, involuntary top–down guidance of attention from working memory. *Journal of Experimental Psychology: Human Perception and Performance*, *31*, 248–261. https://doi.org/10.1037/0096-1523.31.2.248, PubMed: 15826228

Soto, D., Hodsoll, J., Rotshtein, P., & Humphreys, G. W. (2008). Automatic guidance of attention from working memory. *Trends in Cognitive Sciences*, *12*, 342–348. https://doi.org/10.1016/j.tics.2008.05.007, PubMed: 18693131

Stroop, J. R. (1935). Studies of interference in serial verbal reactions. *Journal of Experimental Psychology*, *18*, 643–662. https://doi.org/10.1037/h0054651

Teng, C., Fulvio, J. M., Jiang, J., & Postle, B. R. (2022). Flexible top–down control in the interaction between working memory and perception. *Journal of Vision*, *22*, 3. https://doi.org/10.1167/jov.22.11.3, PubMed: 36205937

Teng, C., & Kravitz, D. J. (2019). Visual working memory directly alters perception. *Nature Human Behaviour*, *3*, 827–836. https://doi.org/10.1038/s41562-019-0640-4, PubMed: 31285620

Teng, C., & Postle, B. R. (2021). Spatial specificity of feature-based interaction between working memory and visual processing. *Journal of Experimental Psychology: Human Perception and Performance*, *47*, 495–507. https://doi.org/10.1037/xhp0000899, PubMed: 33475418

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*, 2022–2037. https://doi.org/10.1037/a0037454, PubMed: 25089574

Whitehead, P. S., Ooi, M. M., Egner, T., & Woldorff, M. G. (2019). Neural dynamics of cognitive control over working memory capture of attention. *Journal of Cognitive Neuroscience*, *31*, 1079–1090. https://doi.org/10.1162/jocn_a_01409, PubMed: 30938591

Wolff, M. J., Ding, J., Myers, N. E., & Stokes, M. G. (2015). Revealing hidden states in visual working memory using electroencephalography. *Frontiers in Systems Neuroscience*, *9*, 123. https://doi.org/10.3389/fnsys.2015.00123, PubMed: 26388748

Wolff, M. J., Jochim, J., Akyürek, E. G., & Stokes, M. G. (2017). Dynamic hidden states underlying working-memory-guided behavior. *Nature Neuroscience*, *20*, 864–871. https://doi.org/10.1038/nn.4546, PubMed: 28414333

Yu, Q., Teng, C., & Postle, B. R. (2020). Different states of priority recruit different neural representations in visual working memory. *PLoS Biology*, *18*, e3000769. https://doi.org/10.1371/journal.pbio.3000769, PubMed: 32598358

Zhang, H., & Luo, H. (2023). Feature-specific reactivations of past information shift current neural encoding thereby mediating serial bias behaviors. *PLoS Biology*, *21*, e3002056. https://doi.org/10.1371/journal.pbio.3002056, PubMed: 36961821

Unconnected Proof

Downloaded from http://direct.mit.edu/jocn/article-pdf/doi/10.1162/jocn_a_02310/2501480/jocn_a_02310.pdf?casa_token=VLDfMCJQ6UAAAAA:63PiRA1U-xfVX1sUK_IgFiFVeddBeHgHOMm8ObNCYPzfPDY8tiN688x0-HCEDB4viY_Go by University of Wisconsin, Madison, Jacqueline Fulvio on 23 April 2024