Cognitive control during scene categorization: The role of identity repetition and timing in congruence sequence effects

*Published Version:*
Tronelli, V., Codispoti, M., De Cesarei, A. (2026). Cognitive control during scene categorization: The role of identity repetition and timing in congruence sequence effects. THE QUARTERLY JOURNAL OF EXPERIMENTAL PSYCHOLOGY, 79(1), 55-68 [10.1177/17470218251335293].

*Availability:*
This version is available at: https://hdl.handle.net/11585/1015994 since: 2025-05-19

*Published:*

DOI: http://doi.org/10.1177/17470218251335293

# Cognitive control during scene categorization:
# The role of identity repetition and timing in Congruence Sequence Effects

Virginia Tronelli, Maurizio Codispoti, Andrea De Cesarei

Department of Psychology, University of Bologna

Keywords: Categorization; Natural Scenes; Cognitive Control; Event Files; Time-Course; Congruence Sequence Effects

Corresponding Author: Andrea De Cesarei
Department of Psychology
Viale Berti Pichat, 5
Bologna (BO)
ITALY
andrea.decesarei@unibo.it

# Abstract

Cognitive control abilities include maintaining goal-directed behaviors in spite of the incongruence between habitual and desired responses. In interference paradigms, slower responses to incongruent compared to congruent trials are observed; this interference is reduced after incongruent trials (Congruence Sequential Effect; CSE), suggesting that the control exerted to counteract interference in the previous trial also propagates into the following trial. Moreover, a larger CSE is observed when trial features are repeated. Binding-retrieval accounts suggest that trial features that occur in the same time frame are bound together in an episodic representation; if a feature is repeated in the next trial, the control state that was active in the previous trial is also reactivated, resulting in a modulation of congruence effects. However, previous studies that used stimulus sets characterized by intra-category variability (e.g., faces; scenes) observed CSE modulation by the repetition of response categories, but were inconclusive concerning whether repeating the identity of a stimulus may modulate CSE. The present study investigates whether episodic stimulus representations include both stimulus identity and response category information, by comparing the impact of the repetition of novel pictures (no identity repetition) and of frequent pictures (in which identity is repeated over trials) in a picture-word interference task. Results indicated that stimulus identity was not critical in the modulation of CSE, and that CSE was little affected by response-stimulus interval. Altogether, the present results contribute to the understanding and theoretical specification of sequential effects.

In everyday activities, as well as in cognitive tasks, the mismatch between novel or weakly established behaviors and well-learned or habitual behaviors leads to conflict that eventually dampens performance. One aspect of cognitive control refers to the ability to maintain stable and goal-directed behaviors, despite interference from mismatching responses (Botvinick et al., 2001; Monsell, 2003; Egner & Hirsch, 2005; Egner, 2023).

In a laboratory setting, tasks that measure cognitive control require participants to respond to task-relevant features while ignoring task-irrelevant features that have habitual response associations and can be either congruent/corresponding or incongruent/non-corresponding with the correct response. A lack of correspondence/congruence leads to a decrease in performance, as in the Simon (Simon & Rudell, 1967), flanker (Eriksen, 1974), Stroop (Stroop, 1935), or Picture Word Interference (PWI; Rosinski, 1977) tasks (Cohen, 2017). In the PWI task, participants name the subject of pictures while ignoring a word that appears together with the picture; the word can be congruent or incongruent with the subject of the pictures (Rosinski, 1977). Results show slower responses to the incongruent trials compared to the congruent ones, reflecting interference provoked by the incongruent words and suggesting similar interference mechanisms despite of the type of to-be-processed material (Van Maanen et al., 2009).

Behavioral interference can be reduced when the preceding trial is incongruent, compared with congruent (Congruence Sequential Effect, CSE; Gratton et al., 1992). Concerning the cognitive processes implied in the CSE, recent models emphasize the role of learning (Egner, 2014; Abrahamse et al., 2016; Dignath et al., 2019; Frings et al., 2020). According to these models, trial characteristics that occur in the same time frame (e.g., in the same trial) are bound together in a short-term episodic representation that can be retrieved in the next trial occurrence. Trial characteristics that can be bound in the episodic representation (event files, Hommel et al., 2004; see also object files in Kahneman, Treisman, & Gibbs, 1992) are perceptual (e.g., Dignath & Kiesel, 2021; Dignath et al., 2019; Grant et al., 2020; De Cesarei et al., 2023; Logan and Zbrodoff, 1979; but see also Kelber et al., 2024, for the binding of more abstract stimulus properties), motor (e.g., Braem, Verguts, Notebaert, 2011; De Cesarei et al., 2023), and/or related to the activation of control states, such as the attentional and processing strategies that have been previously employed (Egner, 2014). Trial-specific event files (Egner, 2014; Hommel et al., 2004; Frings et al., 2020) can then be retrieved when a feature is repeated, resulting in a reactivation of the preceding control state and in a more pronounced CSE compared to trials in which no control state is reactivated (repetition vs. change trials; Dignath et al., 2019). In a recent study (Dignath and Kiesel, 2021), a larger CSE was observed in context repetitions compared to context changes; the contextual feature was the spatial location of the stimulus, which could be at the top or bottom of the monitor.

Previous studies have investigated which stimulus features may modulate CSE if repeated. A stronger CSE was also observed when the format of the target and distractor was repeated from one trial to another, rather than when it was changed (Dignath et al., 2019, Exp. 1: arabic number vs. spelled-out number word for the first experiment; Exp. 2: color patch vs. spelled-out color word for the second experiment). In other studies, the CSE was larger when stimulus modalities (visual or auditory) were repeated compared to when they alternated, provided that they were linked to task set (Grant et al., 2020, 2022), and when the gender of a task-irrelevant voice was repeated across trials, compared with when it changed (Spapé & Hommel, 2008). Overall, these results indicate that high-level stimulus features (i.e., features that are shared with sibling stimuli) are represented in episodic memories, which eventually modulate sequential effects. However, it is less clear whether such episodic memories also represent stimulus-specific features, such as its identity. Recently, De Cesarei et al. (2023) investigated CSE in a task requiring the categorization of natural pictures, and observed a larger CSE when the identity of frequent natural pictures was repeated in two successive trials, compared to when it changed. These results may suggest that the identity of a specific picture can be stored in an episodic representation. However, in this previous study, a change in frequent picture identity also implied a change in picture category, therefore not allowing the issue of whether the episodic representations involve superordinate category stimulus, or specific picture identity, to be disentangled.

Here, we will examine whether stored stimulus representations only include features that are relevant to the task that participants are performing, but that may be shared with other stimuli (e.g., the number of elements; the location; the superordinate category), or also details that are unique to the stimulus that has been shown, such as its identity. To this end, participants will be engaged in a picture-word interference (PWI) task in which they will categorize natural scenes; in order to generalize results to different categorization tasks, participants will be asked to respond to the superordinate category of the foreground object, to the number of foreground elements, or the location of the scene. Each picture will be presented along with a verbal label that can refer either to the same or to the other response category (e.g., an animal picture with the “animal” label is a congruent trial; an indoor picture with the “outdoor” label is an incongruent trial; etc.). In one group of participants, all images from each stimulus category will be the same (e.g., one animal and one vehicle), so that a response repetition from one trial to another will imply the repetition of the same image (identity repetition); in another group of participants, all images will be different, so that the response repetition will only imply the repetition of the same response category (e.g., one animal followed by another animal). If stimulus representation in episodic memory involves the binding and retrieval of specific picture identity, then we expect to find a larger CSE in repetition trials in the group with the same, compared to different, stimuli. If, on the other hand, the episodic representations only involve the response category, then we expect to observe no differences in the modulation of the CSE by repetition between the two groups.

An additional question concerns whether the CSE reflects a sudden and short-lived reactive mechanism, or whether it takes time to activate and lasts in time, reflecting a proactive control system that remains active in preparation for the next trial. For this reason, previous studies investigated the temporal properties (onset and duration) of CSE. A significant CSE in conditions with a response-stimulus interval (RSI) from 500 up to 2000 ms was observed in Stroop and PWI paradigms (Egner et al., 2010; Duthoo et al., 2014; Schiltenwolf et al., 2022), while no significant CSE was observed after a 2250 ms RSI (Egner et al., 2010; Duthoo et al., 2014, Exp. 1). Based on these results, it was suggested

(Egner et al., 2010) that CSE reflects short-lived and phasic enhancements of the attention in reaction to the conflict. Similarly, Hommel & Frings (2020) described the decline over time of event file representations, which are thought to be responsible for CSE (Egner, 2014; Abrahamse et al., 2016; Dignath et al., 2019; Frings et al., 2020). However, Duthoo et al. (2014) modulated the RSI, creating a bias that meant that slower, compared to faster, RSIs had a greater probability of occurring; it was observed that, with such RSI biasing, CSE was also observed at longer time intervals. Also, Schiltenwolf and colleagues (2022) showed that when S-R repetitions were eliminated a stable CSE was observed up to 5000 ms. Taking these results together, CSE may reflect a combination of both reactive and proactive processes: it may reflect the transient-reactive control at faster intervals, and the proactive control may occur slowly at slower intervals. Finally, it has been suggested that it could take time for cognitive control in response to relevant events (e.g., errors, mismatch, conflict, emotional cues) to be activated, and for the stages of triggering inhibitory control of ongoing activities, orienting to the source of the expectancy violation, and reconfiguring the task setting to be set up (Wessel, 2018). Here, we will examine the time course (onset and duration) of CSE, presenting the response-stimulus intervals varying from 750 to 5000 ms. If CSE occurs in fast RSI only (Egner et al., 2010; Duthoo et al., 2014, Exp. 1), it could reflect reactive-transitory processes that vanish with time (Egner et al., 2010; Hommel, 2022). If CSE occurs in all RSIs (Duthoo et al., 2014, Exp. 2), it may reflect proactive mechanisms that persist over time in anticipation of the imminent stimulus (Botvinick et al., 2001). In the latter scenario, If CSE occurs in slow RSI, it may reflect cognitive control that takes time to be activated after relevant stimuli (Wessel, 2018).

## Method

### Participants

A total of 82 participants (53 females, age M = 23.49, SD = 4.03), who were fluent Italian-speaking participants with normal or corrected vision, took part in the experiment. A power analysis was conducted using R (package pwrss; Bulus, 2023) to determine the required sample size based on the four-way interaction of interest (Current Congruency × Previous Congruency × Response Repetition × Picture Frequency), a desired $\eta^2_p = 0.1$ (medium-large effect according to Cohen, 1988; see also Dignath et al. 2019, Grant et al., 2022 and Kelber et al., 2024 for similar partial eta squared values), 80% power, and a low correlation ($r = 0.24$) between differential repeated measures (Incongruent-Congruent differentials, calculated on an independent data sample). The power analysis indicated that the required sample size to observe this interaction effect would be 46 participants. Out of the 82 recruited participants, a total of 40 participants were randomly assigned to the Novel Picture condition while 42 were part of the Frequent Picture condition. Data were analyzed from 80 participants as two participants (both in the Frequent Picture condition) failed to comply with response instructions (one misunderstood the assignment of response keys, and another one disclosed a neurodevelopmental condition, which was an exclusion criterion, after the completion of the experiment). The experimental protocol conformed to the declaration of Helsinki and was approved by the Ethical Committee of the University of Bologna.

## Stimuli

The experimental stimuli comprised 576 different pictures. The picture database was created along three orthogonal dimensions: content (animal or vehicle), number of foreground elements (one or two), and scenario (indoors or outdoors). Hence, each picture portrayed one or two foreground elements (animal or vehicle) in an indoor or outdoor scenario, for a total of eight possible combinations, with equal probability. Pictures were in color and were balanced for brightness and contrast, pixel intensity M = 153 and SD = 5.05 on a 0-255 scale. Each full-screen picture was resized to a 1280 x 1024 pixel monitor subtending 20°30’ horizontal x 16°20’ vertical degrees of visual angle. The distance between the monitor and the participant was 94 cm. Each picture was shown together with a word that was congruent (50%) or incongruent (50%) with the content of the picture (e.g., a picture that represents a dog with the word “vehicle” is an incongruent trial, while the same picture with the word “animal” is a congruent one). The words were printed in Italian, and they were: “veicolo” (“vehicle”), “animale” (“animal”), “uno” (“one”), “due” (“two”), “interno” (“indoors”), and “esterno” (“outdoors”). The white words in Courier New 70 font were displayed at the center of the monitor, surrounded by a black rectangle (15°8’ horizontal x 2°33’ vertical degrees of visual angle).

### Procedure

After filling out the informed consent form, the participant was accompanied to the experimental room. The room illumination was 3 lux, measured using a diode-type digital luxmeter. The task was divided into 3 blocks of 192 trials; the order of the blocks was counterbalanced across participants. Each block involved a different task: superordinate categorization of the content of elements in the forefront (animal or vehicle), enumeration of the number of elements (one or two), or categorization of the scenario the pictures depicted (outdoors or indoors). The participant responded to the picture by ignoring the word and giving equal importance to speed and accuracy, pressing one of two keys (J or N) on the computer keyboard with two fingers of the dominant hand. The block order, as well as the response key, was counterbalanced across participants. The experiment lasted approximately 30 min.

Prior to each experimental block, 8 practice trials, which involved different images from those used in the experimental blocks, were presented in order to let participants familiarize themselves with the task. After each block, the participant was allowed to have a short break.

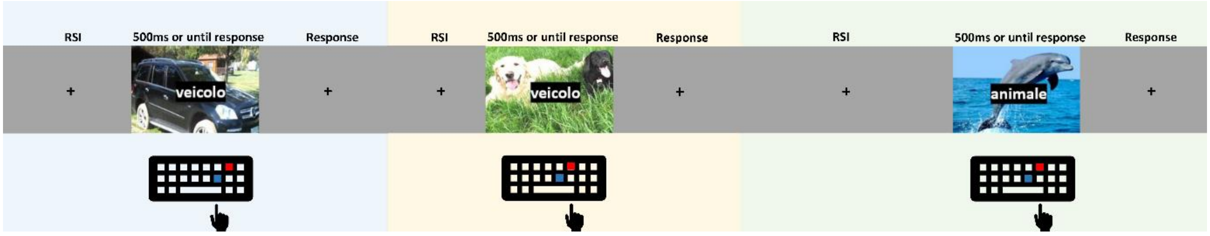
Each trial started with a picture which was presented for 500 ms or until the participant’s response was given. If the participant did not respond while viewing the image, a central fixation cross appeared until they responded. After an RSI ranging from 750 to 5000 ms, the next trial started; an example is reported in Figure 1. The different RSI levels (750, 1000, 1500, 2000, 3000, 5000 ms) were evenly distributed and randomized across trials. Stimuli were presented in pseudo-random sequences that had the following specific constraints: not more than 5 consecutive trials with the same or alternate responses, not more than 5 consecutive trials with the same or alternate congruence conditions, and not more than 5 consecutive trials with the same or alternate RSI levels. Out of the 576 stimuli, half were congruent (C) and half incongruent (I). Considering the congruence level of the previous trial, each of the four possible sequences (PreC-C, PreC-I, PreI-C, PreI-I) was presented with equal probability (48 trials for each congruence condition, in each task block). Moreover, each of these four combinations was equally paired with each of the RSIs resulting in 8 trials for each condition, in each task block. The experiment was run using E-Prime 2.0 Professional.

## Groups

The Picture Word Interference (PWI) task was administered to two groups of participants, which only differed in terms of the pictures used in the task. One group (hereafter, "Novel Picture" condition) viewed each picture of the whole 576 picture dataset, and each picture was seen only once and never repeated. The other group (hereafter, "Frequent Picture" condition) viewed 6 pictures which were selected from the original 576 dataset. Across participants, these 6 pictures were assigned to each different task, i.e., superordinate categorization (animal vs. vehicle), environment categorization (indoor vs. outdoor), enumeration (one vs. two foreground elements). Participants were randomly assigned to experimental groups.

**Figure 1**

*Trial procedure*



## Response Time Analysis

Since response times were not normally distributed, we transformed the reaction times (RTs) using a base-10 logarithm (Whelan, 2010) Practice trials, trials which followed an error or in which an inaccurate response was given, the first trial of each block, and trials with RTs more than 2.5 standard deviations (SD) away from the transformed RT mean were excluded from the RT analyses. Hence, 91% of the total trials was included in the analyses. A Block factor with 3 levels (Content, Numeric, Scenario) was collapsed since the interaction with the Current and Previous Congruence was not significant. Log-transformed RT data were averaged and analyzed with a repeated-measures ANOVA, with the factor Picture Frequency as the between-subjects factor (2 levels: Novel, Frequent) and with the following within-subject factors: RSI (3 levels: Fast [750-1000ms], Medium [1500-2000ms], Slow [3000-5000ms]), Repetition (2 levels: Repeated response, Changed response), Previous Congruence (2 levels: Congruent [PreC], Incongruent [PreI]) and Current Congruence (2 levels: Congruent [C], Incongruent [I]). Huynh–Feldt correction was used when appropriate. The same pattern of results was also observed both when analyzing raw RTs and reciprocal-corrected RTs (Moscoso del Prado Martin, 2008) as an alternate way to correct for non-normality. The partial eta squared statistic ($\eta^2_p$) was calculated and reported.

## Accuracy Analysis

Practice trials and the first trial of each block were excluded from the accuracy analyses. The same ANOVA design used for reaction times was analyzed. ANOVA results are reported in Table 2, while a full description of accuracy results is reported in Supplementary materials along with a descriptive table (S1).

## Data availability

Experimental data and study materials are available at OSF at the link https://osf.io/w3ryb/?view_only=b90d167777394609a868fd5dbbef05bb.

# Results

## Repetition and Congruence

Response times are reported in Figures 2 and 3, and RT descriptive statistics are reported in Table 1. We observed a significant main effect of Current Congruence, F(1, 78) = 97.07, $p < .001$, $\eta^2_p = .55$, with slower responses for incongruent compared with congruent trials (Congruence effect; M = 685.23 ms, SD = 117.99; M = 661.96 ms, SD = 163.48, respectively). This effect was further modulated by a significant interaction between Current and Previous Congruence, F(1, 78) = 13.82, $p < .001$, $\eta^2_p = .15$ with a reduction of the Congruence effect when trials were preceded by incongruent (PreI condition), compared with congruent (PreC condition), trials. Following this interaction, slower responses were observed for incongruent compared with congruent trials both when they were preceded by congruent trials (PreC condition, F(1, 78) = 72.54, $p < .001$, $\eta^2_p = .48$) and when they were preceded by incongruent trials (PreI condition, F(1, 78) = 38.07, $p < .001$, $\eta^2_p = .32$). Moreover, responses to incongruent trials were slower when they were preceded by a congruent compared with an incongruent trial, F(1, 78) = 5.63, p = .020, $\eta^2_p = .07$, while responses to congruent trials were slower when they were preceded by incongruent, compared with congruent, trials, F(1, 78) = 10.72, p = .002, $\eta^2_p$ =.12.

**Table 1.** *RTs in all experimental conditions*

| | | | RSI | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| | | | Fast | | Medium | | Slow | | |
| | | | Actual Congruency | | | | | | |
| | | | congruent | incongruent | congruent | incongruent | congruent | incongruent | CSE |
| Picture Frequency | Response Repetition | Previous Congruency | M (SD) | M (SD) | M (SD) | M (SD) | M (SD) | M (SD) | M (SD) |
| novel | repetition | congruent | 698.09<br>(129.32) | 733.22<br>(145.35) | 689.28<br>(124.91) | 732.96<br>(142.61) | 731.82<br>(141.27) | 782.43<br>(176.14) | 22.66<br>(59.46) |
| | | incongruent | 703.64<br>(125.26) | 717.55<br>(130.17) | 709.15<br>(137.25) | 730.01<br>(151.25) | 760.71<br>(155.15) | 783.40<br>(174.75) | |
| | change | congruent | 709.71<br>(125.92) | 716.16<br>(116.36) | 711.53<br>(129.77) | 744.99<br>(144.41) | 736.48<br>(138.45) | 767.40<br>(163.44) | 2.59<br>(40.82) |
| | | incongruent | 716.69<br>(137.46) | 726.25<br>(123.32) | 714.69<br>(129.50) | 725.55<br>(128.97) | 726.29<br>(128.64) | 762.37<br>(155.03) | |
| frequent | repetition | congruent | 581.08<br>(168.67) | 614.35<br>(212.30) | 575.32<br>(157.10) | 625.76<br>(216.49) | 624.15<br>(185.25) | 653.09<br>(192.16) | 34.072<br>(62.48) |
| | | incongruent | 588.61<br>(196.43) | 584.26<br>(177.09) | 610.60<br>(222.22) | 607.10<br>3(206.31) | 639.84<br>(187.80) | 660.85<br>(203.86) | |
| | change | congruent | 612.63<br>(175.46) | 619.02<br>(191.93) | 594.21<br>(173.66) | 621.09<br>(205.49) | 607.65<br>(171.71) | 644.79<br>(200.47) | 1.15<br>(46.15) |
| | | incongruent | 612.75<br>(205.50) | 627.70<br>(209.28) | 607.15<br>(200.98) | 618.95<br>(194.95) | 617.53<br>(175.12) | 647.44<br>(206.36) | |

**Table 2.** *Results for repeated-measure ANOVAs for Log-transformed Reaction Times (RTs) and Accuracy*

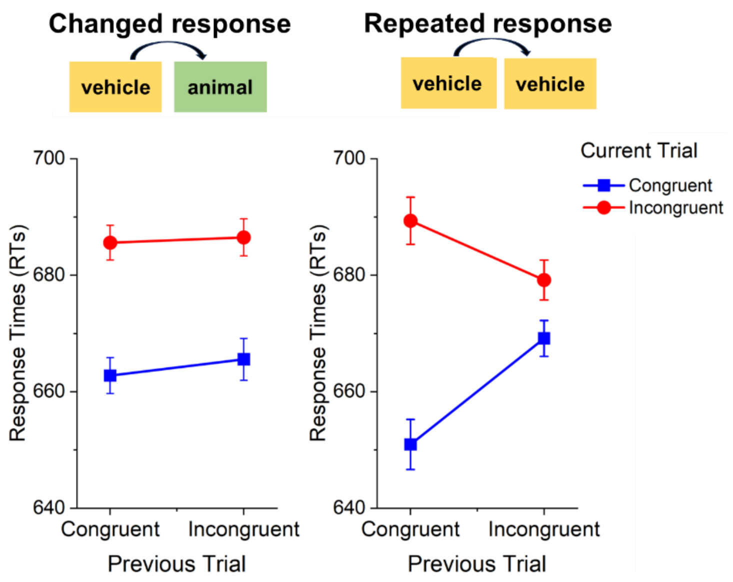
| | | Log10(RTs) | | | Accuracy | | |
|---|---|---|---|---|---|---|---|
| Source | df | F | p | $\eta^2_p$ | F | p | $\eta^2_p$ |
| C | 1,78 | 97.07 | <.001 | .55 | 32.41 | <.001 | .29 |
| C X Pic | 1,78 | .83 | .365 | .01 | 1.79 | .184 | .02 |
| C X Rep | 1,78 | 2.22 | .140 | .03 | .01 | .902 | .00 |
| C X RSI | 2,156 | 2.93 | .058 | .04 | 1.26 | .286 | .02 |
| PreC | 1,78 | 1.41 | .239 | .02 | 10.90 | .001 | .12 |
| PreC X C | 1,78 | 13.82 | <.001 | .15 | 4.76 | .032 | .06 |
| PreC X Pic | 1,78 | .33 | .566 | .00 | .99 | .322 | .01 |
| PreC X Rep | 1,78 | 1.39 | .243 | .02 | .02 | .889 | .00 |
| PreC X RSI | 2,156 | 2.83 | .062 | .03 | 3.04 | .050 | .04 |
| Rep | 1,78 | 4.62 | .035 | .06 | .49 | .486 | .01 |
| Rep X Pic | 1,78 | 1.65 | .203 | .02 | .60 | .439 | .01 |
| Rep X RSI | 2,156 | 31.62 | <.001 | .29 | .33 | .721 | .00 |
| Rep X C X Pic | 1,78 | 1.08 | .303 | .01 | .00 | .969 | .00 |
| Rep X PreC X C | 1,78 | 6.99 | .010 | .08 | .45 | .502 | .01 |
| Rep X PreC X Pic | 1,78 | 1.79 | .185 | .02 | .11 | .746 | .00 |
| Pic | 1,78 | 15.80 | <.001 | .17 | 4.80 | .031 | .06 |
| Pic X RSI | 2,156 | .17 | .766 | .00 | 1.28 | .281 | .02 |
| Pic X PreC X C | 1,78 | .00 | .974 | .00 | 3.64 | .060 | .04 |
| Pic X Rep X PreC X C | 1,78 | .58 | .447 | .01 | 2 .09 | .152 | .03 |
| RSI | 2,156 | 45.55 | <.001 | .37 | 2.07 | .130 | .03 |
| RSI X C X Pic | 2,156 | .13 | .875 | .00 | .08 | .927 | .00 |
| RSI X PreC X Pic | 2,156 | 1.35 | .262 | .02 | 1.71 | .185 | .02 |
| RSI X Rep X C | 2,156 | .58 | .559 | .01 | 2.59 | .078 | .03 |
| RSI X Rep X Pic | 2,156 | 6.55 | .002 | .08 | 3.49 | .033 | .04 |
| RSI X Rep X PreC | 2,156 | 4.54 | .012 | .05 | 1.79 | .172 | .02 |
| RSI X Rep X C X PreC | 2,156 | .33 | .716 | .00 | .44 | .627 | .01 |
| RSI X PreC X C | 2,156 | 2.55 | .081 | .03 | .88 | .415 | .01 |
| RSI X PreC X C X Pic | 2,156 | .22 | .803 | .00 | 1.18 | .309 | .01 |
| RSI X Rep X PreC X Pic | 2,156 | .14 | .867 | .00 | .01 | .993 | .00 |
| RSI X Rep X C X Pic | 2,156 | .36 | .696 | .00 | 3.02 | .052 | .04 |
| RSI X Rep X C X PreC X Pic | 2,156 | 1.21 | .300 | .01 | 1.62 | .204 | .02 |

*Note*. Pic=Picture Frequency; RSI= response-stimulus interval; Rep=Repetition; PreC= Previous Congruence; C=Current Congruence.

Response repetition had a significant main effect on response times, F(1, 78) = 4.62, p = .035, $\eta^2_p$ = .06, with slower responses for changed response (M = 675.08 ms, SD = 169.25) compared with repeated response (M = 672.08 ms, SD = 172.80). Moreover, response Repetition interacted significantly with Previous and Current Congruence, three-way interaction F(1, 78) = 6.99, p = .010, $\eta^2_p$ = .08 (Figure 2). Following this result, the interaction between Previous and Current Congruence was assessed in the repeated response condition and the changed response condition, and it was significant only for repeated responses, F(1, 78) = 18.35, p < .001, $\eta^2_p$ = .19, but not for changed ones, F(1, 78) = .41, p = .525, $\eta^2_p$ < .001, with a reduction of the Congruence effect when trials were preceded by incongruent (PreI condition), compared with congruent (PreC condition), trials in the Repeated condition. Following the interaction between Previous and Current congruence in repeated conditions, significantly slower RTs were observed for incongruent versus congruent trials preceded by congruent trials, F(1, 78) = 56.09, p < .001, $\eta^2_p$ = .41, and when preceded by incongruent ones, F(1, 78) = 10.92, p = .001, $\eta^2_p$ = .12. Slower responses to congruent trials were observed when they were preceded by congruent, compared with incongruent, trials, F(1, 78) = 16.58, p < .001, $\eta^2_p$ = .17, and for incongruent trials when they were preceded by incongruent, compared with congruent, trials, F(1, 78) = 5.69, p = .019, $\eta^2_p$ = .07. Focusing on the significant three-way interaction between Repetition, Previous, and Current Congruence, we examined whether Repetition mostly affects Congruence effects for trials preceded by congruent, or incongruent, conditions. A significant Repetition x Congruence effect was only observed in the PreC: F(1, 78) = 6.47, p = .013, $\eta^2_p$ = .08. Following this interaction, faster responses in PreC-C trials in the repeated response condition compared to the changed response condition were observed, F(1, 78) = 14.31, p < .001, $\eta^2_p$ =.15.

**Figure 2**

*Mean RT data as a function of Previous and Current Congruence in Repeated and Changed responses*

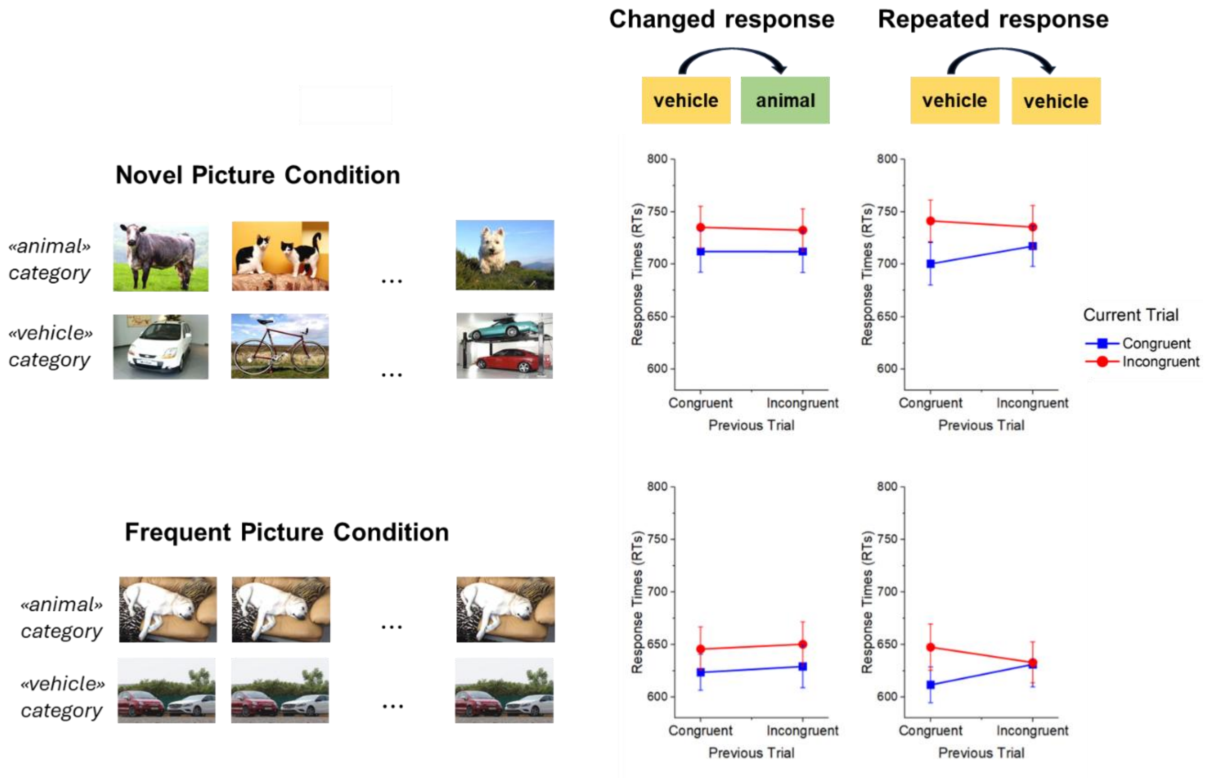


*Note.* Error bars reflect ±1 within-subject standard errors of the mean (Cousineau, 2005).

A main effect of Picture Frequency was observed with faster response times in the Frequent Picture condition than in the Novel Picture condition, F(1, 78) = 15.80, p < .001, $\eta^2_p$ = .17 (M = 616.50 ms; SD = 187.61; M = 730.56 ms; SD = 130.10 respectively). Finally, no significant four-way interaction between Picture condition, Repetition, Previous Congruence, and Current Congruence was observed, F(1, 78) = .58, p = .447, $\eta^2_p$ = .01 (Figure 3)[1]. In order to corroborate the result of the Picture Set x Repetition x Previous Congruency x Current Congruency interaction, we conducted an equivalence analysis (Lakens, 2017) to determine whether the difference between the CSEs in repeated and changed responses could be considered statistically equivalent in the two experimental groups (Novel Picture and Frequent Picture). Based on an equivalence test, we rejected the presence of effects more extreme than -0.7 to 0.7 Cohen's dz (medium-large effect), t(78) = 2.08, p = 0.021. Finally, an additional Bayesian ANOVA confirmed that evidence for an interaction involving Picture Set was against its inclusion in the model, either as a four-way interaction between Picture Set x Repetition x Previous Congruency x Current Congruency, $BF_{excl}$ = 1.50, or as a three-way interaction between Picture Set x Previous Congruency x Current Congruency, $BF_{excl} = 4.91$.

**Figure 3**

*Mean RT data as a function of Previous and Current Congruence in Repeated and Changed responses, separated by Novel and Frequent picture conditions*



*Note.* Error bars reflect ±1 within-subject standard errors of the mean (Cousineau, 2005).

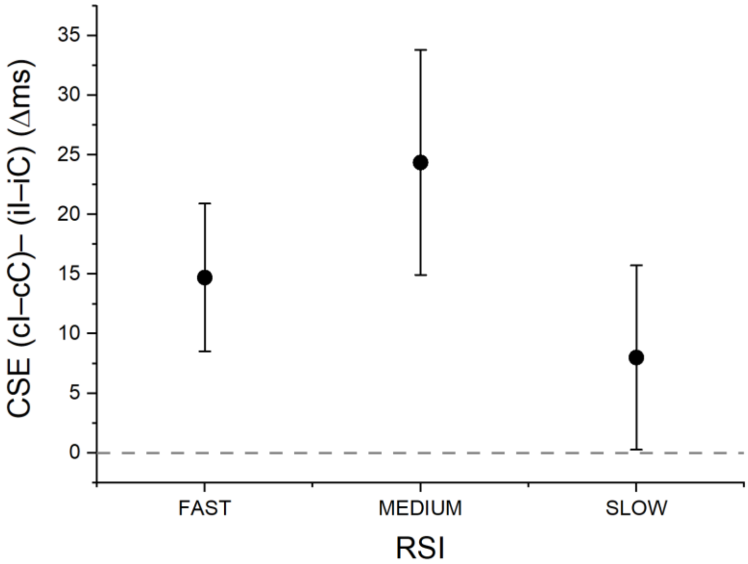
## Response Stimulus Interval (RSI)

Response-Stimulus interval (RSI) significantly modulated response times, F(2, 156) = 45.55, $p < .001$, $\eta^2_p = .37$, with slower RTs for the long RSI (3000-5000 ms, M = 696.48 ms, SD = 175.85) compared with the medium RSI (1500-2000 ms, M = 664.64 ms, SD = 173.17), F(1, 78) = 92.06, $p < .001$, $\eta^2_p = .54$, and no difference between the medium and the fast RSI (750-1000 ms, M = 660.17 ms, SD = 168.11), F(1, 78) = .39, p = .535, $\eta^2_p < .001$.

The three-way interaction between RSI, Previous, and Current Congruence was not significant, F(2, 156) = 2.55, $p = .081$, $\eta^2_p = .03$ (Figure 4). Focusing on the difference between CSE at each RSI level, no differences were observed between any RSI level (fast vs. medium, F(1, 78) = 1.01, $p = .317$, $\eta^2_p = .01$; medium vs. slow, F(1, 78)= 3.28, $p = .074$, $\eta^2_p = .04$; fast vs. slow, F(1, 78) = 1.01, $p = .318$, $\eta^2_p = .01$). However, when testing each RSI level separately, a significant CSE was only observed only at fast and medium RSI levels, F(1, 78) = 6.52, $p = .013$, $\eta^2_p = .08$ and F(1, 78) = 9.48, $p = .003$, $\eta^2_p = .11$ respectively, but not for the slowest RSI level (3000-5000 ms), F(1, 78) = .63, $p = .431$, $\eta^2_p = .01$.

**Figure 4**

*Congruence sequence effect at each RSI level*



*Note.* Error bars reflect ±1 within-subject standard errors of the mean (Cousineau, 2005).

A significant interaction was observed between RSI and Repetition, F(2, 156) = 31.62, $p < .001$, $\eta^2_p = .29$. Following the interaction, the main effect of Repetition was significant (Fs ≥ 7.52, ps ≤ .008, $\eta^2_p s \geq .09$) across all RSI intervals. Responses were faster for Repeated conditions compared to Changed conditions in the Fast and Medium RSI intervals (M = 653.28 ms, SD = 169.88; M = 661.84 ms, SD = 175.84, respectively, for Repeated; M = 667.42 ms, SD = 168.00, M = 667.63 ms, SD = 171.65, respectively, for

Changed). Conversely, in the Slow RSI interval, responses were slower for Repeated conditions (M = 702.71 ms, SD = 180.78) compared to Changed conditions (M = 690.15 ms, SD = 173.45).

The three-way interaction between RSI, Repetition, and Picture Frequency was also significant, F(2, 156) = 6.55, p = .002, $\eta^2_p$ =.08. Following this result, the interaction between RSI and Repetition was assessed for each Picture Frequency condition. The interaction was significant in the Novel Picture condition, F(2, 78) = 4.95, p = .010, $\eta^2_p$ =.11, however no significant effect of repetition was observed in any RSI level in the Novel picture Condition, Fs ≥ 3.14, ps ≤ .084, $\eta^2_p$s ≥ .07. The interaction between RSI and Repetition was also significant in the Frequent Picture condition, F(2, 78) = 26.58, p < .001, $\eta^2_p$ = .40. Following this interaction, the main effect of Repetition was significant in all RSIs (Fs ≥ 4.55, ps ≤ .039, $\eta^2_p$s ≥ .10) in the Frequent Picture condition. Responses in fast and medium RSI levels were faster in the repeated response condition (M = 592.88 ms, SD = 186.53; M = 605.90 ms, SD = 196.21) than in the changed response condition (M = 617.33 ms, SD = 192.52; M = 611.23 ms, SD = 191.30). In the slow RSI (3000-5000 ms), responses were faster for changed responses (M = 631.12 ms, SD = 186.76) compared with repeated ones (M = 642.96 ms, SD = 188.67).

The three-way interaction between RSI, Repetition, and Previous Congruence was significant, F(2, 156) = 4.54, p = .012, $\eta^2_p$ =.05. Following this result, the interaction between Repetition and Previous Congruence was significant in slow RSI only, F(1, 79) = 6.53, p = .013, $\eta^2_p$ =.08. Analyzing this interaction, the main effect of Previous Congruence was significant in the repeated response condition only, F(1, 78) = 10.80, p = .002, $\eta^2_p$ = .12, with faster responses in the trials following the PreC trials (M = 695.52 ms, SD = 179.21) compared to the PreI trials (M = 710.43 ms, SD = 185.44).

## Discussion

The present study aimed to investigate cognitive control in a picture-word interference task, with a focus on understanding whether stimulus identity adds to category/response repetition in shielding from incongruency, and on determining the temporal profile of congruence sequence effects. Here, performance was modulated by semantic incongruency between the picture and the word, and the effect was more pronounced when preceded by congruent, rather than incongruent, trials (CSE effect; Gratton et al., 1992).

We observed an interference effect (i.e., lower accuracy and slower response times in incongruent trials compared to congruent trials) and a CSE pattern (i.e., a reduction in the differences between incongruent and congruent trials after incongruent trials compared to congruent trials) on both accuracy and response times. Incongruence between the picture and the word dampened performance on both accuracy and RTs, excluding a speed-accuracy trade off. Regarding response time, we also observed a modulation of CSE depending on whether the response category in the actual and previous trials was the same (Repetition condition) or changed (Changed condition). Specifically, we observed a significant CSE in trials involving response repetition, but not in trials involving response change, suggesting a modulation of CSE induced by repetition. This result is consistent with previous studies, in which CSE was diminished or absent in trials involving feature change compared to trials involving feature repetition (De Cesarei et al., 2023; Dignath et al., 2019; Dignath & Kiesel, 2021; Grant et al., 2020; Egner, 2014). According to Spapé & Hommel (2008), previous trial congruence affects the degree of interference only in repeated trials. In particular, they suggest that in the episodic representations, the integration of stimuli, actions, and task-specific information occurs, so that later reactivation of some elements of a representation will also retrieve other elements of the previous trial, including control states that facilitate performance (Egner, 2014; Abrahamse et al., 2016; Dignath et al., 2019; Frings et al., 2020). Here, a CSE was observed in trials involving repetition, but not in trials involving change. Interestingly, consistently with previous results (De Cesarei et al., 2023; Dignath et al., 2021), the change in CSE pattern between repeated and changed conditions depended on faster responses to congruent preceded by congruent stimuli in repeated compared with changed trials, rather than on a modulation in the response to incongruent preceded by congruent trials, or on a modulation in trials preceded by incongruent conditions. This pattern of results suggests that repetition, in addition to interference-contrasting processes (e.g., Spapé & Hommel, 2008; Egner et al., 2014), also relies on facilitatory processes that may speed up responses to congruent stimuli once engaged by retrieval (see also Dunaway & Weismann, 2024). Facilitation has been observed both in classical Stroop paradigms (MacLeod, 1991), and in PWI tasks, where distractors that are semantically related to the target produce facilitation rather than interference (Glaser & Dungelhoff, 1984; Costa et al., 2003; Hantsch et al., 2012). As a consequence, it is possible that cognitive control processes act both on the interference-related and on the facilitation-related component of such tasks.

Which features can be bound together in episodic representations, and under which conditions, is an open question that the present study addressed. Most notably, the present study asked whether target stimulus identity would add to the effects of response category repetition in modulating CSE. This was not the case, and similar CSE were observed both when stimulus identity was repeated and when stimulus identity changed. In the present paradigm, participants either performed in a categorization task that involved the viewing of repeated pictures (Frequent Picture condition), therefore allowing to learn the unique individual features of each target picture, or performed in a categorization task which involved the viewing of always-novel pictures, in which no picture was ever repeated (Novel Picture condition) and no picture-defining perceptual feature could therefore be learnt. If learning of single-picture features helps to form episodic representations of the stimuli, then one would expect more pronounced effects of repetition in the frequent than in the novel condition. Stimulus identity did not add to the effects of repetition, suggesting that response category, but not necessarily scene identity, is represented in episodic memories, and its repetition is a sufficient condition for the observation of a CSE without further modulation by identity. According to this possibility, stimulus identity might have a less pronounced (if any) role, compared with superordinate or more abstract features, in the coding structure which is maintained in the episodic event file. However, several previous studies in the attention, categorization and visual search literature argue against a fixed and hardwired use of information, and promote instead a flexible and diagnostic use of visual information (Di Lollo et al., 2001; Wolfe et al., 2021; De Cesarei et al., 2019; Schyns, 1998; Oliva & Schyns, 2000).

An alternative explanation for the present results is that task requirements promoted cognitive control at the goal-related, rather than at the stimulus-specific, level. Several previous studies (e.g., Grant et al. 2020, 2022; Kelber, 2024) asked whether changes in stimulus features affect CSE independently of the task that is being carried out, or only when they concern task-defining features. These studies indicate that changes in stimulus features only reduce CSE if they are associated with changes in task set. More so, not only perceptual but also abstract features, such as cross-modal intensity, can modulate CSE if associated to a task set (Kelber et al., 2024). Altogether, these results are consistent with a task-set account (Grant et al. 2020, 2022; Kelber, 2024; Hazeltine, et al., 2011), which posits that the features of task-irrelevant stimuli do not trigger control adjustments unless they are associated with task sets or predict upcoming task-relevant information. In PWI tasks, as in several interference tasks (e.g., the Stroop or the flanker task), targets and interfering stimuli share the same task set, either in terms of being the same digit misplaced (and therefore mapping to the same response: flanker task, Eriksen, 1974), or in terms of mapping to the same semantic concept (Stroop and PWI task: Stroop, 1935; Rosinski, 1977); for this reason, targets and distractors in such tasks have been classified as competing “candidate targets”, eventually resulting in response interference within the same task (Buetti, Lleras, & Moore, 2014). In the present PWI task, while word meaning is a distractor that competes with the picture at the level of semantic access (candidate target, or task-relevant distractor), stimulus identity is not associated with task requirements, and is therefore unrelated to the task goal. If this is the case, then other tasks that strongly rely on individual stimulus representation, e.g., an old/new task in which old and new pictures are presented along with a congruent or incongruent old/new label, may be more sensitive to stimulus identity.

Consistently with previous studies, we observed a CSE when the response category was repeated, but not when it changed. In the present paradigm however, repeating a response also involved repeating the category. Therefore, the observed results may be related either to a repetition/change in to-be-categorized stimulus features (e.g., animal or vehicle category, or one or two foreground elements), or to a response repetition/change. According to the binding-retrieval CSE model (Hommel et al., 2004; Frings et al., 2020), stimulus and response features (e.g., actions, task-specific information, etc.) are integrated into episodic representations, that can be retrieved if one feature is repeated. Hence, future studies may adopt confound-minimized designs (which were not employed here to favor the manageability of the task) that control the alternation of category and responses across trials, in order to eliminate repetition of category/response combinations (Braem et al., 2019; Schmidt & Liefooghe, 2016; Weismann et al., 2014) and to dissociate the separate contributions of category and response repetition.

The second question concerned the time course (onset and duration) of CSE. The present study showed similar CSE at least for fast and medium (up to 2000 ms) RSIs. Although the interaction did not reach standard significance (p = .081) and no difference between CSE at any RSI level was observed, no significant CSE was observed at the slowest RSI, corresponding to 3000-5000 ms. While the pattern of results suggests maintaining a cautious approach in interpreting these data, the present results indicate that CSE is little affected by RSI, i.e., shows little decay, and a significant CSE is observed until as late as 2000 ms after the response. At the same time, CSE was already present at the fastest RSI (750-1000 ms) and the present results do not support, for the examined paradigm, the prediction from the error processing literature that proactive cognitive control would have taken time to activate (e.g., Wessel, 2018). The present temporal profile of CSE is consistent with a previous study (Schiltenwolf et al., 2022), in which a constant CSE was observed up to 5000 ms in a conflict task that minimized repetition effects. On the other hand, other studies (e.g., Duthoo et al., 2014; Egner et al., 2010; see also Hommel & Frings, 2020) observed a decline in CSE over time. Future studies might further investigate this issue, possibly taking into additional consideration the differences between the type of material (e.g., natural pictures vs. faces vs. numbers or letters), and comparing tasks that favor proactive vs. reactive cognitive control.

In the present study, we investigated the role of trial features (picture identity vs. response category) and trial timing on the modulation of CSE. It was observed that CSE was modulated by the repetition of response categories, but not of stimulus identity. These results are consistent with previous literature indicating that stimulus features that are not bound to task set are not involved in the modulation of CSE. However, several questions remain open to future research, e.g., concerning individual differences in cognitive control and their cognitive and neural basis. First, previous studies tried to disentangle the effects of experimental manipulations on the proper binding (i.e., what happens at the time of stimulus encounter) and on retrieval (for reviews see Hommel, 2022; Frings, 2020). Moreover, the conditions which promote the binding or retrieval of certain features should be better understood (Hommel, 2022). One such factor, which might modulate either the binding or retrieval of stimulus features, is task set. While here task set promoted superordinate categorization, and we did not observe CSE modulation by stimulus identity, other tasks explicitly asking participants to respond to stimulus identity may reveal that, when it is made task relevant, stimulus identity is coded into event files (e.g., in an old/new task). Together with the result of this and previous studies (e.g., Grant et al. 2020, 2022; Kelber, 2024; Hazeltine, et al., 2011; Egner, 2023), this would support the primary role of task set in the modulation of CSE and, at a broader level, contribute to a better understanding of the conditions which promote cognitive control. Finally, individual differences in nonclinical and clinical populations might be considered. While previous studies indicated that CSE was not modulated by working memory capacity (Meier & Kane, 2012), other individual differences might modulate cognitive control. Of note, schizophrenia patients demonstrate deficits in tasks that assess cognitive control, episodic memory and working memory (Barch & Sheffield, 2017). Specifically, Barch and Sheffield (2017) suggest that a decrease in proactive control may impair the ability to represent, maintain, and use goal-relevant information to guide thoughts and actions. One electrocortical measure that is involved in cognitive control and is altered in patients with schizophrenia (Basar-Eroglu et al., 2008) is the amplitude of theta oscillations. Theta oscillations (4-7 Hz) reflect cortical mechanisms summoned by cognitive processes (Cavanagh and Cohen, 2022; Poldrack, 2011) that are involved during conflict conditions compared to non-conflict situations (Cohen, 2014; Cavanagh et al. 2012; Cohen and Cavanagh et al. 2011; Cohen and Donner, 2013; Hanslmayer et al. 2008; Nigbur et al., 2011; 2012; Gratton, 2018), and represent an ideal candidate for the investigation of this facet of cognitive control in nonclinical and clinical population.

## Conclusions

The present study examined the memory mechanisms underlying cognitive control, focusing on the congruency sequence effect in a PWI task. Using natural scenes, similar CSE was observed across different tasks, and showed little decay within the observed interval. Moreover, the present results indicate that response category, but not necessarily stimulus identity, may be bound together in an episodic representation and contribute to the modulation of CSE if retrieved. Altogether, the present results are informative concerning the type and duration of learned representations that are active during an interference task.

# Footnotes

1. The present research design was replicated, with Picture Set manipulated within participants, in a follow-up EEG study (N = 49). In the replication study, a main effect of Current Congruence was observed, F(1, 48) = 16.89, p < .001, $\eta^2_p$ = .26, which was further modulated by Previous Congruence, F(1, 48) = 20.31, p < .001, $\eta^2_p$ = .30, with a result pattern consistent with the present study. As in the present study, a significant three-way interaction between Response Repetition, Previous Congruence, and Current Congruence was observed, F(1, 48) = 6.22, p = .016, $\eta^2_p$ = .12, with significant CSE pattern for repeated responses, F(1, 48) = 19.41, p < .001, $\eta^2_p$ = .29, but not for trials where responses changed, F(1, 48) = .59, p = .448, $\eta^2_p$ = .01. Consistent with the present data, the four-way interaction between Picture Frequency, Response Repetition, Previous Congruence, and Current Congruence did not reach significance, F(1, 48)=.010, p=.920, $\eta^2_p$ <.001 (novel pictures: repeated response, CSE M = 25.26, SD = 47.00; changed response, CSE M = 7.78, SD = 50.77; frequent pictures: repeated response, CSE M = 29.40, SD = 58.11; changed response, CSE M = 13.85, SD = 46.67).

## Declaration of Conflicting Interests

The authors have no competing interests to declare.

## Funding Acknowledgements

This research received no specific grant from any funding agency in the public, commercial, or not-for-profit sectors.

## References

Abrahamse, E., Braem, S., Notebaert, W., & Verguts, T. (2016). Grounding cognitive control in associative learning. *Psychological Bulletin, 142*(7), 693. https://doi.org/10.1037/bul0000047

Barch, D. M., & Sheffield, J. M. (2017). Cognitive control in schizophrenia: Psychological and neural mechanisms. *The Wiley handbook of cognitive control*, 556-580. https://doi.org/10.1002/9781118920497.ch31

Basar-Eroglu, C., Schmiedt-Fehr, C., Marbach, S., Brand, A., & Mathes, B. (2008). Altered oscillatory alpha and theta networks in schizophrenia. *Brain research*, 1235, 143-152. https://doi.org/10.1016/j.brainres.2008.06.114

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. http://dx.doi.org/10.1037/0033-295X.108.3.624

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

Bulus, M. (2023). pwrss: Statistical Power and Sample Size Calculation Tools. R package version 0.3.1, https://CRAN.R-project.org/package=pwrss.

Buetti, S., Lleras, A., & Moore, C. M. (2014). The flanker effect does not reflect *the processing of “task-irrelevant” stimuli: Evidence from inattentional blindness. Psychonomic Bulletin & Review*, 21, 1231–1237. https://doi.org/10.3758/s13423-014-0602-9

Cavanagh, J. F., & Cohen, M. X. (2022) Frontal Midline Theta as a Model Specimen of Cortical Theta. In P. A. Gable, M. W. Miller & E. M. Bernat (Eds.), *The Oxford Handbook of EEG Frequency* (online ed, pp. 178-201). Oxford Academic. https://doi.org/10.1093/oxfordhb/9780192898340.013.9

Cavanagh, J. F., Zambrano-Vazquez, L., & Allen, J. J. (2012). Theta lingua franca: A common mid-frontal substrate for action monitoring processes. *Psychophysiology, 49*(2), 220-238. https://doi.org/10.1111/j.1469-8986.2011.01293.x

Cohen, J. (1988). Statistical Power Analysis for the Behavioral Sciences (2nd ed.). Hillsdale, NJ:Erlbaum.

Cohen, J. D. (2017). Cognitive control: Core constructs and current considerations. In T. Egner (Ed.), *The Wiley handbook of cognitive control*, (1st ed, pp. 1-28). Wiley Online Library. https://doi.org/10.1002/9781118920497.ch1

Cohen, M. X. (2014). A neural microcircuit for cognitive conflict detection and signaling. *Trends in neurosciences, 37*(9), 480-490. https://doi.org/10.1016/j.tins.2014.06.004

Cohen, M. X., & Cavanagh, J. F. (2011). Single-trial regression elucidates the role of prefrontal theta oscillations in response conflict. *Frontiers in psychology*, 2, 30. https://doi.org/10.3389/fpsyg.2011.00030

Cohen, M. X., & Donner, T. H. (2013). Midfrontal conflict-related theta-band power reflects neural oscillations that predict behavior. *Journal of neurophysiology, 110*(12), 2752-2763. https://doi.org/10.1152/jn.00479.2013

Costa, A., Mahon, B., Savova, V., & Caramazza, A. (2003). Level of categorisation effect: A novel effect in the picture-word interference paradigm. *Language and cognitive processes*, *18*(2), 205-234. https://doi.org/10.1080/01690960143000524

Cousineau, D. (2005). Confidence intervals in within-subject designs: A simpler solution to Loftus and Masson’s method. *Tutorials in quantitative methods for psychology, 1*(1), 42-45. https://doi.org/10.20982/tqmp.01.1.p042

De Cesarei, A., Cavicchi, S., Micucci, A., & Codispoti, M. (2019). Categorization Goals Modulate the Use of Natural Scene Statistics. Journal of Cognitive Neuroscience. 10.1162/jocn_a_01333

De Cesarei, A., D'Ascenzo, S., Nicoletti, R., & Codispoti, M. (2023). Novelty and learning in cognitive control: evidence from the Simon task. *Psychological Research*, 1-17. https://doi.org/10.1007/s00426-023-01813-z

Di Lollo, V., Kawahara, J.., Zuvic, S. M., & Visser, T. A. W. (2001). The preattentive emperor has no clothes: A dynamic redressing. Journal of Experimental Psychology: General, 130(3), 479–492. https:// https://doi.org/10.1037/0096-3445.130.3.479.

Dignath, D., & Kiesel, A. (2021). Further evidence for the binding and retrieval of control-states from the flanker task. *Experimental Psychology.* https://doi.org/10.1027/1618-3169/a000529

Dignath, D., Johannsen, L., Hommel, B., & Kiesel, A. (2019). Reconciling cognitive-control and episodic-retrieval accounts of sequential conflict modulation: Binding of control-states into event-files. *Journal of Experimental Psychology: Human Perception and Performance, 45(*9), 1265. https://doi.org/10.1037/xhp0000673

Dunaway, M. G., & Weissman, D. H. (2024). Both congruent and incongruent trials drive the congruency sequence effect: Novel support for an episodic retrieval view of adaptive control in the prime–probe task.. Journal of Experimental Psychology: Learning, Memory, and Cognition. Advance online publication. https:// dx.doi.org/10.1037/xlm0001349

Duthoo, W., Abrahamse, E. L., Braem, S., & Notebaert, W. (2014). Going, going, gone? Proactive control prevents the congruency sequence effect from rapid decay. *Psychological Research, 78*(4), 483-493. https://doi.org/10.1007/s00426-013-0498-4

Egner, T. (2014). Creatures of habit (and control): a multi-level learning perspective on the modulation of congruency effects. *Frontiers in psychology, 5*, 1247. https://doi.org/10.3389/fpsyg.2014.01247

Egner, T. (2023). Principles of cognitive control over task focus and task switching. *Nature Reviews Psychology*, 1-13. https://doi.org/10.1038/s44159-023-00234-4

Egner, T., & Hirsch, J. (2005). Cognitive control mechanisms resolve conflict through cortical amplification of task-relevant information. *Nature neuroscience, 8*(12), 1784-1790. https://doi.org/10.1038/nn1594

Egner, T., Ely, S., & Grinband, J. (2010). Going, going, gone: characterizing the time-course of congruency sequence effects. *Frontiers in psychology,* 1, 154. https://doi.org/10.3389/fpsyg.2010.00154

Eriksen, B. A., & Eriksen, C. W. (1974). Effects of noise letters upon the identification of a target letter in a non-search task. *Perception & psychophysics*, *16*(1), 143-149. https://doi.org/10.3758/BF03203267

Frings, C., Hommel, B., Koch, I., Rothermund, K., Dignath, D., Giesen, C., ... & Philipp, A. (2020). Binding and retrieval in action control (BRAC). *Trends in Cognitive Sciences, 24*(5), 375-387. https://doi.org/10.1016/j.tics.2020.02.004

Glaser, W. R., & Düngelhoff, F.-J. (1984). The time course of picture-word interference. *Journal of Experimental Psychology: Human Perception and Performance, 10*(5), 640–654. https://doi.org/10.1037/0096-1523.10.5.640

Grant, L. D., Cookson, S. L., & Weissman, D. H. (2020). Task sets serve as boundaries for the congruency sequence effect. *Journal of Experimental psychology: Human perception and performance, 46*(8), 798. https://doi.org/10.1037/xhp0000750

Grant, L.D., Cerpa, S.R., Weissman, D.H. (2022). Rethinking attentional reset: Task sets determine the boundaries of adaptive control. Quarterly Journal of Experimental Psychology 75(6): 1171–1185. doi:10.1177/17470218211047424.

Gratton, G., Coles, M. G., & Donchin, E. (1992). Optimizing the use of information: strategic control of activation of responses. *Journal of Experimental Psychology: General, 121*(4), 480. https://doi.org/10.1037/0096-3445.121.4.480

Gratton, G., Cooper, P., Fabiani, M., Carter, C. S., & Karayanidis, F. (2018). Dynamics of cognitive control: Theoretical bases, paradigms, and a view for the future. *Psychophysiology, 55*(3), e13016. https://doi.org/10.1111/psyp.13016

Hanslmayr, S., Pastötter, B., Bäuml, K. H., Gruber, S., Wimber, M., & Klimesch, W. (2008). The electrophysiological dynamics of interference during the Stroop task. *Journal of cognitive neuroscience, 20*(2), 215-225. https://doi.org/10.1162/jocn.2008.20020

Hantsch, A., Jescheniak, J. D., & Mädebach, A. (2012). Naming and categorizing objects: task differences modulate the polarity of semantic effects in the picture-word interference paradigm. *Memory & cognition, 40*(5), 760-768. https://doi.org/10.3758/s13421-012-0184-6

Hazeltine, E., Lightman, E., Schwarb, H., & Schumacher, E. H.(2011). The boundaries of sequential modulations: Evidence for set-level control. *Journal of Experimental Psychology: Human Perception and Performance*, 37, 1898-1914. https://doi.org/10.1037/a0024662

Hommel, B. (2022). The Control of Event-File Management. *Journal of Cognition, 5*(1), 1-15. https://doi.org/10.5334/joc.187

Hommel, B., & Frings, C. (2020). The disintegration of event files over time: Decay or interference? *Psychonomic Bulletin & Review*, 27, 751-757. https://doi.org/10.3758/s13423-020-01738-3

Hommel, B., Proctor,R.W., & Vu,K.P. (2004). A feature-integration account of sequential effects in the Simon task. *Psychological research.* 68, 1-17. https://doi.org/10.1007/s00426-003-0132-y

Kahneman, D., Treisman, A., & Gibbs, B.J. (1992). The reviewing of object files: Object-specific integration of information. Cognitive Psychology 24(2), 175-219

Kelber, P., Mackenzie, I. G., & Mittelstädt, V. (2024). Cognitive control in cross-modal contexts: Abstract feature transitions of task-related but not task-unrelated stimuli modulate the congruency sequence effect. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 50*(6), 902-919. https://doi.org/10.1037/xlm0001300

Lakens, D. (2017). Equivalence Tests: A Practical Primer for t Tests, Correlations, and Meta-Analyses. Social Psychological and Personality Science, 8(4), 355-362. https://doi.org/10.1177/1948550617697177

MacLeod, C. M. (1991). Half a century of research on the Stroop effect: an integrative review. *Psychological bulletin, 109*(2), 163. https://doi.org/10.1037/0033-2909.109.2.163

Meier, M.E., Kane, M.J. (2012). Working memory capacity and Stroop interference: global versus local indices of executive control. Journal of Experimental psychology. Learning, Memory, and Cognition. 39(3):748-759. DOI: 10.1037/a0029200.

Monsell, S. (2003). Task switching. *Trends in Cognitive Sciences, 7*(3), 134-140. https://doi.org/10.1016/S1364-6613(03)00028-7

Moscoso Del Prado Martín, F. (2010). *A Theory of Reaction Time Distributions*. Available at: http://cogprints.org/6310/1/recinormal.pdf.

Nigbur, R., Cohen, M. X., Ridderinkhof, K. R., & Stürmer, B. (2012). Theta dynamics reveal domain-specific control over stimulus and response conflict. *Journal of Cognitive Neuroscience, 24*(5), 1264-1274. https://doi.org/10.1162/jocn_a_00128

Nigbur, R., Ivanova, G., & Stürmer, B. (2011). Theta power as a marker for cognitive interference. *Clinical Neurophysiology, 122*(11), 2185-2194. https://doi.org/10.1016/j.clinph.2011.03.030

Oliva, A., & Schyns, P.G. (2000). Diagnostic Colors Mediate Scene Recognition. Cognitive Psychology 41(2), 176-210

Poldrack, R. A. (2011). Inferring mental states from neuroimaging data: from reverse inference to large-scale decoding. *Neuron, 72*(5), 692-697. https://doi.org/10.1016/j.neuron.2011.11.001

Rosinski, R. R. (1977). Picture-word interference is semantically based. *Child Development,* 643-647. https://doi.org/10.2307/1128667

Schiltenwolf, M., Kiesel, A., & Dignath, D. (2022). No temporal decay of cognitive control in the congruency sequence effect. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *49*(8), 1247. https://doi.org/10.1037/xlm0001159

Schmidt, J.R., Liefooghe, B. (2016). Feature Integration and Task Switching: Diminished Switch Costs after Controlling for Stimulus, Response, and Cue Repetitions. PLoS ONE 11(3): e0151188. doi:10.1371/journal.pone.0151188

Schyns, P. G. (1998). Diagnostic recognition: Task constraints, object information, and their interactions. Cognition, 67, 147–179.

Simon, J. R., & Rudell, A. P. (1967). Auditory S-R compatibility: The effect of an irrelevant cue on information processing. *Journal of Applied Psychology*, 51, 300-304. https://doi.org/10.1037/h0020586

Spapé, M. M., & Hommel, B. (2008). He said, she said: Episodic retrieval induces conflict adaptation in an auditory Stroop task. *Psychonomic Bulletin & Review*, 15, 1117-1121. https://doi.org/10.3758/PBR.15.6.1117

Stroop, J. R. (1935). Studies of interference in serial verbal reactions. *Journal of experimental psychology, 18*(6), 643. https://doi.org/10.1037/h0054651 van Maanen, L., van Rijn, H., & Borst, J. P. (2009). Stroop and picture—word interference are two sides of the same coin. *Psychonomic bulletin & review*, 16, 987-999. https://doi.org/10.3758/PBR.16.6.987

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

Wessel, J. R. (2018). An adaptive orienting theory of error processing. *Psychophysiology, 55*(3), e13041. https://doi.org/10.1111/psyp.13041

Whelan, R. (2010). Effective analysis of reaction time data. The Psychological Record, 58(3): 475–482.

Wolfe, J.M. (2021). Guided Search 6.0: An updated model of visual search. Psychonomic Bulletin and Review, 28(4):1060-1092. doi: 10.3758/s13423-020-01859-9

# Author Note

Contribution of each author to the present work, according to the Contributor Roles Taxonomy (CRediT), was as follows: Conceptualization: VT, MC, ADC; Data curation: VT, ADC; Formal analysis: VT, ADC; Funding acquisition: MC, ADC; Investigation: VT; Methodology: VT, MC, ADC; Project administration: MC, ADC; Resources: MC, ADC; Software: VT, MC, ADC; Supervision: MC, ADC; Validation: VT, MC, ADC; Visualization: VT, MC, ADC; Writing—original draft: VT, MC, ADC; Writing—review and editing: VT, MC, ADC.