Cognitive Neurodynamics (2025) 19:114
https://doi.org/10.1007/s11571-025-10307-0

RESEARCH ARTICLE

Reward masks the learning of cognitive control demand

Bettina Bustos[1,2] · **Jiefeng Jiang**[1,2] · **Wouter Kool**[3]

Received: 15 October 2024 / Revised: 5 June 2025 / Accepted: 22 June 2025

Abstract
Cognitive control refers to a set of cognitive functions that modulate other cognitive processes to align with internal goals. Recent research has shown that cognitive control can flexibly adapt to internal and external factors such as reward, effort, and environmental demands. This suggests that learning processes track changes in these factors and drive an optimization process to determine how cognitive control should be applied in changing situations. In real life, multiple factors often simultaneously affect how cognitive control is deployed. However, previous studies mainly concern how cognitive control adjusts to changes in a single factor. Here, we investigate how cognitive control learns to adjust to two concurrently changing factors: statistical regularity in cognitive control demand and performance-contingent reward. We consider two competing hypotheses: reward promotes cognitive control to adjust to cognitive control demand, and the processing of reward information obstructs the adaptation to cognitive control demand. In our experiment, statistical regularity in cognitive control demand is manipulated within subjects such that some stimuli require higher levels of cognitive control than others. Reward is manipulated across subjects. Using a computational model that captures temporal changes in cognitive control, we find that in the absence of reward, participants can adjust to different levels of cognitive control demand. Importantly, when performance-contingent reward is available, participants fail to adapt to changes in cognitive control demand. The findings support the hypothesis that reward blocks the learning of cognitive control.

Keywords Cognitive control · Reinforcement · Item-specific proportion congruence · Motivation · Reward

Introduction

Our ability for cognitive control enables us to reconfigure mental functions to implement effortful, non-routine, and goal-directed behavior (Botvinick et al. 2001; Egner 2017; Miller and Cohen 2001). A key feature of cognitive control is its flexibility to adapt to changing environmental and internal factors (Jiang et al. 2014, 2015; Monsell 2003). Such flexibility is thought to optimize the tradeoff (Ritz et al. 2022; Shenhav et al. 2013) between potential gains, such as reward (Botvinick and Braver 2015), goal achievement (Devine and Otto 2022; Devine et al. 2024), opportunity costs (Otto and Daw 2019), and potential costs such as cognitive effort (Kool and Botvinick 2018; Kool et al. 2010; Shenhav et al. 2017; Westbrook et al. 2013).

These factors are typically studied in isolation. For example, humans adjust cognitive control in response to demands posed by the environment. In the classic Stroop task (Stroop 1935), they increase their use of cognitive control to the proportion of demanding incongruent trials (e.g., the word GREEN printed in red; Braem et al. 2019; Gratton et al. 1992; Jacoby et al. 2003). Such adaptation does not only happen across temporal contexts (Egner 2007; Schmidt 2013; Spinelli et al. 2022; Ullsperger et al. 2005) but also in response to statistical contingencies between items and demand (Bugg and Dey 2018; Bugg et al. 2011; Chiu et al. 2017; Jiang et al. 2020a). Computationally, this flexibility can be captured by the reinforcement-learning framework (Jiang et al. 2014, 2015), which posits that the brain

✉ Jiefeng Jiang
jiefeng-jiang@uiowa.edu

✉ Wouter Kool
wkool@wustl.edu

[1] Department of Psychological & Brain Sciences, University of Iowa, Iowa City, IA 52242, USA

[2] Cognitive Control Collaborative, University of Iowa, PBSB 373, 340 Iowa Avenue, Iowa City, IA 52242, USA

[3] Department of Psychological & Brain Sciences, Washington University in St. Louis, CB 1125, One Brookings Drive, St. Louis, MO 63130, USA

Published online: 18 July 2025

learns to predict future demand based on a running average of experienced congruency, and adjusts cognitive control accordingly.

Cognitive control also flexibly responds to available reward, with evidence showing that performance-contingent reward increases accuracy and reduces response times by enhancing attentional engagement and cognitive control (Balleine and Dickinson 1998; Botvinick and Braver 2015; Chaillou et al. 2017; Chiew and Braver 2014; Fröber and Dreisbach 2014).

In real life, cognitive control needs to flexibly adapt to a host of factors, including reward and predictable statistical structures, simultaneously. For example, rewards may be given for performing a challenging task that affords the learning of cognitive control demand. Recent work suggests that such rewards may bias cognitive control strategies by promoting proactive control but at the cost of decreased flexibility (Hefer and Dreisbach 2016, 2017, 2020) or increasing reliance on congruent distractors (Fröber and Lerche 2023). However, little is known about how these factors jointly affect flexible cognitive control, and more specifically, the ability to update learned control settings in response to changing environmental demands.

We consider two hypotheses of how cognitive control simultaneously adapts to changes in performance-related incentives and statistical contingencies predicting the demand for cognitive control. On the one hand, performance-contingent rewards may increase the motivation to tailor the attentional state to the structure of cognitive control demand. On the other hand, the increased motivation to engage cognitive control triggered by reward may interfere with the learning of updated demand contingencies. For example, because both reward and learning of statistical contingency produce prediction errors (PEs, the discrepancy between the predicted and actual outcome), these may interfere and lead to reduced learning of statistical contingency in the presence of reward. Moreover, structure learning has been characterized as carrying a control cost (Collins 2017), thus increased cognitive control on a focal task may reduce availability for control applied to updating statistical structure.

To adjudicate between these hypotheses, we manipulated the statistical structure of cognitive control demands by changing the item-specific proportion congruence (ISPC) in a picture-word Stroop task. In an initial 'Inducer' phase of the experiment, certain stimuli mostly appeared in congruent trials (i.e., having low cognitive control demand), whereas others mostly appeared in incongruent trials (i.e., having high cognitive control demand). Humans are sensitive to these differences between trial types, adopting a more focused attentional state for items that are mostly incongruent and a more relaxed one for items that are mostly congruent (Bugg and Crump 2012; Bugg and Dey 2018; Bugg et al. 2011; Crump et al. 2006; Logan and Zbrodoff 1979; Suh and Bugg 2021). In a second, Diagnostic, phase this statistical structure was altered, with all items appearing with 50% percent congruence. If participants are sensitive to this change in the statistical structure of the task, item-specific cognitive control settings should regress from their previously focused and relaxed states to identical and intermediate states for all stimuli. To test whether reward and structure learning interact, one group of participants received performance-contingent reward in the Diagnostic phase, whereas the other group did not. We analyzed the data using conventional analyses of error rates and reaction time (RT) data and computational modeling using a reinforcement-learning framework (Chiu et al. 2017; Sutton and Barto 2018). To preview the results, we find that the presence of performance-contingent reward reduces the learning of the ISPC, thus supporting the hypothesis that reward interferes with the learning of statistical contingency of cognitive control. To our knowledge, this work provides the first evidence of the interaction between reward and the learning of cognitive control demand in driving adaptive behavior.

Method

Participants

No-reward condition. One hundred and twenty-nine participants from Washington University in St. Louis provided informed consent and earned class credit for participation. Twenty-one participants were removed for not meeting the accuracy threshold of 80%, resulting in a final sample of 108 participants (Age $M = 19.4$, $SD = 1.09$, 83 female, 23 male, 2 preferred not to answer). All procedures were approved by Washington University in St. Louis Institutional Review Board.

Reward condition. One hundred and twenty-two participants were recruited on Amazon Mechanical Turk to participate in the experiment and provided informed consent. Six participants were removed for not meeting the accuracy threshold of 80%, resulting in a final sample of 116 participants (Age $M = 40.66$, $SD = 11.85$, 49 female, 67 male). All procedures were approved by Washington University in St. Louis Institutional Review Board.

Stimuli

The stimuli used in both conditions of this study were a subset of a larger set of stimuli from Bugg and colleagues (2011). The stimulus set consisted of sixteen line drawings

of four birds, four cats, four fish, and four dogs, along with four capitalized words (BIRD, CAT, FISH, DOG). On each trial, a word (45px) was superimposed on a line drawing (215 × 161px). Both the word and the line drawing were presented at the center of the screen.

Procedure

Participants completed the Reward condition or the No-reward condition. All procedures were held constant across both conditions in the Inducer phase of the experiment with no reward. The only difference between the two conditions occurred in the Diagnostic phase, namely that in the Reward condition participants could earn performance-contingent monetary reward, whereas participants in the No-reward condition did not.

Each trial started with the presentation of a picture-word Stroop stimulus until a valid response was made or until the response deadline of 3000ms was met (Fig. 1A). Participants were instructed to respond to the animal that appeared in the picture-word Stroop stimuli using previously learned response keys. They were also instructed to ignore the animal names superimposed on the picture. On congruent trials, the identity of the picture and the word matched (e.g., a picture of a cat with the word CAT superimposed), whereas on incongruent trials, the identity of the picture and the word conflicted (e.g., a picture of a cat with the word DOG superimposed). Stimulus-response mappings (e.g., the correspondence between keys on the keyboard [A, S, D, or F] and animal pictures [dog, fish, cat, bird]) were randomized across participants. Trials were separated by an inter-trial interval (ITI) of 1000ms. A crosshair was presented at the center of the screen during the ITI. The crosshair was replaced with a reward stimulus on trials in which participants received reward (only in the Diagnostic phase of the Reward condition).

After the instructions, participants first completed the Inducer phase, which consisted of three blocks of 240 trials each for a total of 720 trials. Specifically, trials with two types of animal pictures were 90% congruent (mostly congruent, or MC condition, Fig. 1B), and trials with the other two types of animal pictures were congruent 10% of the time (mostly incongruent, or MI condition, Fig. 1B). The animal-congruence mappings were randomized across participants.

After the Inducer phase, participants in both conditions completed a Diagnostic phase consisting of one block of 192 trials. In this phase, participants performed the same task, except that the previously learned ISPC biases were changed– all stimuli had the same 50% probability of appearing in a congruent trial. Note that in the Diagnostic phase, the proportion congruence condition (MC or MI) of a stimulus refers to its previous ISPC in the Inducer phase. Critically, participants in the Reward condition were instructed that quick and accurate responses would be probabilistically rewarded (Fig. 1A). However, whether a trial was rewarded was determined randomly. We did not use RTs for assigning reward to avoid congruent trials, which are typically faster than incongruent trials, to be more frequently rewarded than incongruent trials and hence introduce confounds. We do not expect this procedure to affect our conclusions, as our

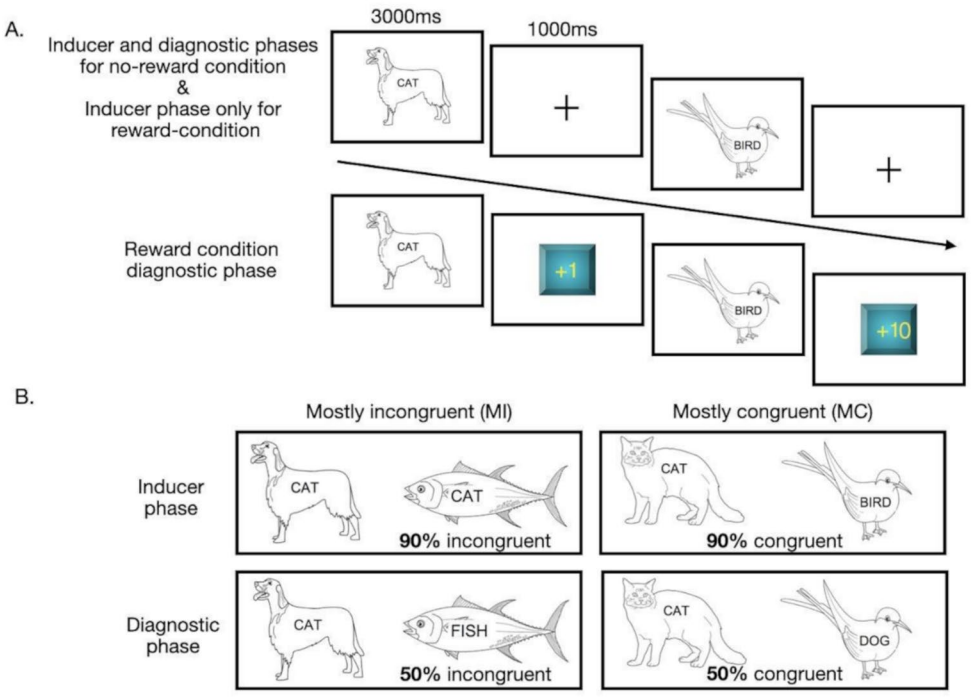

Fig. 1 Experimental design. (**A**) Trial sequence. Exemplar trials for Diagnostic and Inducer phases for each condition. (**B**) ISPC for each phase. ISPC is the same in both conditions. MI and MC categories are counterbalanced across participants

main focus is the global effect of reward rather than the specific reward policy.

Specifically, one MI animal category and one MC animal category carried a 90% chance of paying out high reward (10 points) for accurate responses while the other two animal categories had a 10% chance of paying out low reward (1 point). On trials in which they received reward, participants were presented with either "+1" or "+10" points displayed within a hexagonal teal-colored shape during the ITI inter-trial interval (1000 ms). In the No-reward condition, the procedure of the Diagnostic phase remained identical to the Reward condition, with the exception that no reward was provided.

Before starting the main task, participants went through a practice phase, in which they were extensively instructed on (1) the correct stimulus-response mappings, and (2) the picture-word Stroop task. Participants practiced the main task for 16 trials. To prevent the practice phase from confounding the learning in the main task, the practice phase used different line drawings (two per category) than those used in the main task. The practice phase is explained in detail in the Supplemental Materials.

Conventional statistical analysis

RTs faster than 200ms or slower than 3,000ms were excluded, consistent with prior research using this task (e.g., Bugg and Dey 2018; Bugg et al. 2011). In addition, only correct responses were included in the analysis of RTs. Both exclusions eliminated less than 1% of the trials. Mean RTs and error rates for each phase and each condition are presented in Table 1. We conducted 2 × 2 repeated-measures analyses of variance (rmANOVAs) with factors of ISPC (MC and MI) and Trial Type (Congruent and Incongruent) for both error rates and RT in the Induction phase as well as the Diagnostic phase. Additionally, to examine the dynamics of ISPC learning in the Diagnostic phase, we split the Diagnostic phase into two halves and repeated the aforementioned rmANOVAs on each half. Finally, we examined whether the amount of reward obtained in the Diagnostic phase of the Reward condition had any effect on ISPC using independent sample *t*-tests. Here, the ISPC effect was quantified as the interaction between PC and Trial Type using the difference in congruency effect (i.e., incongruent - congruent) between MC and MI conditions (Bugg et al. 2011). All statistical analyses were conducted using Pingouin (version number: 0.3.12; Vallat (2018).

Model-based analysis

Modeling the learning of cognitive control demand. We used a reinforcement-learning model to more precisely capture the dynamics of how participants tracked the statistical structure of cognitive control demand (i.e., ISPC in this study) throughout the different phases. The model predicts the probability of encountering an incongruent trial for each animal category across the task. A model was applied to each category to account for the manipulation of PC at the animal category level. The predicted incongruency of the presented category was modeled and updated on each trial as $P_i \leftarrow P_i + \alpha\,(c - P_i)$, where P_i is the predicted incongruency (i.e., the predicted probability of encountering an incongruent trial) for animal category i presented on the current trial, and c encodes the incongruency at the

Table 1 Mean Reaction Time (ms) and Error Rates

Phase	Condition	Item type PC	DV	Trial type		Congruency effect
				Con	Inc	
Induction	No-reward	MC	RT	750 (117)	860 (159)	110 (81)
			Error rate	4.4% (0.03)	7.3% (0.06)	2.9% (0.06)
		MI	RT	753 (129)	806 (140)	53 (60)
			Error rate	4.1% (0.04)	5.9% (0.04)	1.7% (0.04)
	Reward	MC	RT	801 (171)	914 (202)	113 (84)
			Error rate	4.1% (0.04)	6.2% (0.07)	2.1% (0.06)
		MI	RT	816 (183)	860 (186)	44 (60)
			Error rate	4.3% (0.05)	6.0% (0.05)	1.5% (0.05)
Diagnostic	No-reward	MC	RT	770 (160)	836 (177)	66 (76)
			Error rate	5.5% (0.05)	7.5% (0.07)	2.0% (0.05)
		MI	RT	771 (166)	813 (170)	42 (73)
			Error rate	5.4% (0.06)	6.6% (0.6)	1.2% (0.05)
	Reward	MC	RT	749 (159)	835 (189)	86 (75)
			Error rate	4.5% (0.07)	6.7% (0.08)	2.2% (0.06)
		MI	RT	756 (159)	817 (171)	61 (62)
			Error rate	4.1% (0.05)	5.8% (0.06)	1.7% (0.05)

Note. Values in parentheses indicate standard deviation. PC = proportion congruence; rt = reaction time; mc = mostly congruent; mi = mostly incongruent; dv = dependent variable. Note that items in the diagnostic phase were 50% congruent

current trial, where congruent and incongruent trials were coded as 0 and 1, respectively. In other words, on each trial, the updated P_i(i.e., on the left side of the equation) is a sum of the current value of P_i(i.e., on the right side of the equation) and an updating term based on the current-trial PE, defined as $(c - P_i)$, weighted by a learning rate α. For the Inducer phase, all P_is were initialized as 0.5 to reflect a neutral prediction of incongruency. The initial P_i of the Diagnostic phase was set to the final P_i of the Inducer phase to simulate the retention of ISPC. The model incorporates two learning rates (α s), one for the Inducer phase and one for the Diagnostic phase.

Behavioral model fitting. To model how P(subscripts are removed to indicate a vector of all trial-wise values for the variable) influences RT, a linear model using $|PE|$ as a predictor was constructed to predict trial-wise RT (Chiu et al. 2017; Jiang et al. 2020a; Muhle-Karbe et al. 2018). The absolute value of PE is used because only the magnitude but not the direction of PE is required to examine the strength of learning. Specifically, a larger $|PE|$ indicates more deviation of the predicted probability of incongruency from the actual (in)congruency, and when used to guide cognitive control, will cause suboptimal processing and slower RT (Jiang et al. 2014, 2015). In addition to the model-based regressors, we included five binary predictor variables. The first coded for the congruency of the current trial, c, to account for the congruency effect, and the four resulting binary predictor variables coded for each of the four animal categories present on the current trial to account for potential biases that may differentially influence RTs for each of the animal categories. All trials were included in the trial-by-trial modeling of P and $|PE|$, but only correct trials were included in the multiple linear regression predicting RT.

The learning rates are the only free parameters of this model. To determine the best-fitting α s, we implemented a standard minimization function (using the L-BFGS-B algorithm) that was run with 30 different randomly selected starting points for each participant to avoid local minima. For each participant, the parameter fits of the model with the smallest sum of squared errors (SSE) of predicted RT over all trials was selected. Within each phase, the learning rate was shared by all category-specific learning models. The optimization procedure also produced the regression coefficient of|PE| (β_{PE}), which indicates how much RT scales with the magnitude of prediction error.

Model-based statistical analysis. Because β_{PE} was generated based on the best-fitting learning model, it quantifies the strength of learning: As discussed above, a positive β_{PE} indicates that RT slows with PE magnitude and thus suggests the behavior is consistent with the learning model (Jiang et al. 2015). In other words, a positive coefficient suggests that the participant learned the current ISPC and reacted to deviations from it by slowing their response. To probe the difference in ISPC learning between conditions, we conducted an independent-samples *t*-test on β_{PE} between conditions (Reward or No-reward) in the Inducer and Diagnostic phases separately. The same analysis was also performed on learning rates. Welsh's *t*-test was used to account for unequal sample sizes between the Reward and No-reward conditions.

Results

Because we more mostly interested in the dynamics of learning captured by the reinforcement-learning model, we will briefly report the results of the conventional statistical analysis below. The full results are in Supplementary Notes 1 and 2. We report the descriptive statistics of RTs and error rates across trial types, phases and conditions in Table 1. Their statistical results are listed in Table 2 (RTs) and 3 (error rates). Note that the ANOVA results of the No-reward condition are originally reported in Bustos and colleagues (2024).

RT data from the Inducer phase showed that participants in both conditions learned the ISPC (No-reward condition: $p < .001$, $\eta_p^2 = 0.26$; Reward condition: $p < .001$, $\eta_p^2 = 0.43$ (Fig. 2). They further suggested reduced learning of the new (unbiased) ISPC in the Reward condition compared to the No-reward condition during the second half of the Diagnostic phase, resulting in a significant difference in the change in ISPC across the first and second halves of the phase between conditions ($p = .03$, $\eta_p^2 = 0.01$). We found no difference in the ISPC effect between the low- and high-reward items in the Diagnostic phase of the Reward condition. Additionally, we did not observe a significant correlation between the behavioral ISPC effect and the amount of reward accumulated ($r = -.04$, $p = .66$). This allowed us to rule out the possibility of reward amount being a confounding factor of the results (Supplementary Note 3). In summary, reward seems to interfere with learning of the ISPC regularities.

To evaluate the robustness of our findings and address concerns about potential sample differences, we conducted a three-way mixed-effects model examining the effect of Reward, PC, Congruency and their interactions on RT in the diagnostic phase. This model showed no main effect of reward on response time ($p = .72$), indicating that simply receiving a reward did not globally alter response speed. Second, the model indicated that reward modulated the congruency effect ($p = .022$). This aligns with previous findings (e.g., Prével et al. 2021) showing that reward can bias cognitive control strategies, such as enhancing proactive control when incongruent trials are rewarded.

Table 2 Conventional statistical results: RT

Phase	Condition	Effect	DOF	F value	*P* value	Effect size η_p^2
Inducer	No-reward	Main effect of trial type	(1,107)	258.53	< 0.001	0.71
	Reward		(1,115)	191.86	< 0.001	0.63
	No-reward	Main effect of ISPC	(1,107)	18.85	< 0.001	0.15
	Reward		(1,115)	8.79	< 0.001	0.07
	No-reward	Trial type × ISPC interaction	(1,107)	36.63	< 0.001	0.26
	Reward		(1,115)	87.80	< 0.001	0.43
Diagnostic	No-reward	Main effect of trial type	(1,107)	114.37	< 0.001	0.52
	Reward		(1,115)	174.84	< 0.001	0.61
	No-reward	Main effect of ISPC	(1,107)	12.15	0.146	0.02
	Reward		(1,115)	0.86	0.36	0.01
	No-reward	Trial type × ISPC interaction	(1,107)	5.57	0.02	0.05
	Reward		(1,115)	15.75	< 0.001	0.12
Diagnostic: First half	No-reward	Main effect of trial type	(1,107)	71.05	< 0.001	0.40
	Reward		(1,115)	126.51	< 0.001	0.52
	No-reward	Main effect of ISPC	(1,107)	0.35	0.56	0.003
	Reward		(1,115)	0.84	0.36	0.01
	No-reward	Trial type × ISPC interaction	(1,107)	10.54	0.002	0.09
	Reward		(1,115)	8.56	0.004	0.07
Diagnostic: Second half	No-reward	Main effect of trial type	(1,107)	54.21	< 0.001	0.34
	Reward		(1,115)	133.8	< 0.001	0.54
	No-reward	Main effect of ISPC	(1,107)	3.05	0.08	0.03
	Reward		(1,115)	0.49	0.48	0.004
	No-reward	Trial type × ISPC interaction	(1,107)	0.03	0.86	< 0.001
	Reward		(1,115)	8.51	0.004	0.07

Note. Items in the diagnostic phase were 50% congruent. dof = degree of freedom

Table 3 Conventional statistical results: Error rate

Phase	Condition	Effect	DOF	F value	*P* value	Effect size η_p^2
Inducer	No-reward	Main effect of trial type	(1,107)	38.61	< 0.001	0.26
	Reward		(1,115)	22.28	< 0.001	0.16
	No-reward	Main effect of ISPC	(1,107)	6.14	0.015	0.05
	Reward		(1,115)	0.11	0.74	< 0.001
	No-reward	Trial type × ISPC interaction	(1,107)	4.26	0.04	0.04
	Reward		(1,115)	0.99	0.32	0.01
Diagnostic	No-reward	Main effect of trial type	(1,107)	21.02	< 0.001	0.16
	Reward		(1,115)	28.34	< 0.001	0.20
	No-reward	Main effect of ISPC	(1,107)	2.1	0.15	0.02
	Reward		(1,115)	1.71	0.74	0.01
	No-reward	Trial type × ISPC interaction	(1,107)	1.53	0.22	0.01
	Reward		(1,115)	0.54	0.46	0.004
Diagnostic: First half	No-reward	Main effect of trial type	(1,107)	16.61	< 0.001	0.13
	Reward		(1,115)	10.70	0.001	0.09
	No-reward	Main effect of ISPC	(1,107)	0.35	0.55	0.003
	Reward		(1,115)	0.72	0.40	0.001
	No-reward	Trial type × ISPC interaction	(1,107)	2.13	0.15	0.02
	Reward		(1,115)	3.15	0.08	0.03
Diagnostic: Second half	No-reward	Main effect of trial type	(1,107)	5.14	0.03	0.06
	Reward		(1,115)	28.35	< 0.001	0.20
	No-reward	Main effect of ISPC	(1,107)	2.08	0.15	0.02
	Reward		(1,115)	1.71	0.74	0.01
	No-reward	Trial type × ISPC interaction	(1,107)	0.05	0.83	0.02
	Reward		(1,115)	0.54	0.46	0.004

Note. Items in the diagnostic phase were 50% congruent. dof = degree of freedom

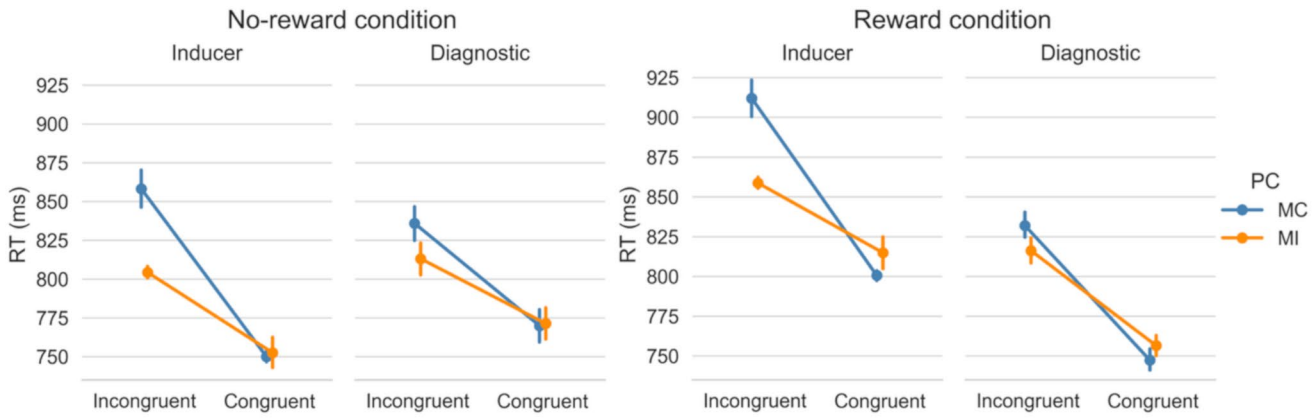

Fig. 2 RT results. (**A**) RTs by trial type (congruent or incongruent) and proportion congruency (mostly congruent, MC, or mostly incongruent, MI) in the No-reward condition for each phase. (**B**) RTs by trial type (congruent or incongruent) and proportion congruency (mostly congruent, MC, or mostly incongruent, MI) in the Reward condition for each phase

Table 4 Three-way mixed-effects model: Reward × Congruency × Phase (Inducer vs. Transfer)

Predictor	Estimate (β\beta)	SE	z-value	*p*-value
Intercept	805.599	13.978	57.634	<0.001
Reward	56.580	19.424	2.913	0.004
Congruency	-59.240	2.193	-27.012	<0.001
Phase	12.193	3.380	3.607	<0.001
Reward × Congruency	-4.016	3.048	-1.318	0.188
Reward × Phase	-51.633	4.697	-10.994	<0.001
Congruency × Phase	6.240	4.780	1.306	0.192
Reward × Congruency × Phase	-16.213	6.642	-2.441	0.015

To account for potential group differences, we conducted an additional three-way mixed-effects model examining the effect of Reward, Congruency, Phase (Inducer vs. Transfer) and their interactions on RT (see Table 4). Since the Inducer phase serves as a baseline in which neither group received reward, this approach controlled for sample-related variability before the introduction of reward. This model showed a small but reliable effect indicating that the effect of reward on congruency differed between phases (p=.015, f^2=0.013). This result supports the interpretation that the observed effects in the Transfer phase were driven by how reward influenced the adaptation of control settings over time, rather than inherent group differences.

Briefly, the error rate data showed two patterns: (1) higher error rates in incongruent than congruent trials and (2) no significant difference in ISPC effects between the reward conditions, which indicates that the difference in ISPC effects in RT data was unlikely to be attributable to a speed-accuracy tradeoff.

In summary, reward seems to interfere with learning of the ISPC. Below, we provide support for this hypothesis using model-based measures that offer a more nuanced account of learning dynamics.

Model-based results

Inducer phase. We used trial-wise model estimates of unsigned PE of incongruency to predict RT in a linear model. Previous studies have shown that trial-wise RT scales with PE in an ISPC design (Chiu et al. 2017; Jiang et al. 2020a). Replicating these findings, we found that β_{PE} was significantly above zero in both conditions (No-reward condition: $t(107)=3.50$, $p<.001$, $d=0.34$; Reward condition: $t(115)=10.7$, $p<.001$, $d=0.99$, Fig. 3A). As a validation, we tested whether the regression weight of PE was correlated with the ISPC effect in RT. In both conditions, the behavioral ISPC effect was strongly associated with the regression weight of $|PE|$(No-reward: $r=.59$, $p<.001$, Reward: $r=.63$, $p<.001$).

Importantly, the average regression weight of $|PE|$(i.e., β_{PE}) was significantly smaller in the No-reward condition than the Reward condition, $t(211)=-2.63$, $p=.01$, $d=0.35$ (No-reward: $M=54$, SE=13.90, Reward: $M=101$, SE=11.45), suggesting that the Reward group learned the ISPC better than the No-reward group. To account for this between-group difference, the β_{PE} of the Inducer phase was used as a baseline in the Diagnostic phase analysis (reported below). In addition to β_{PE}, we also compared learning rates between the two conditions. To account for

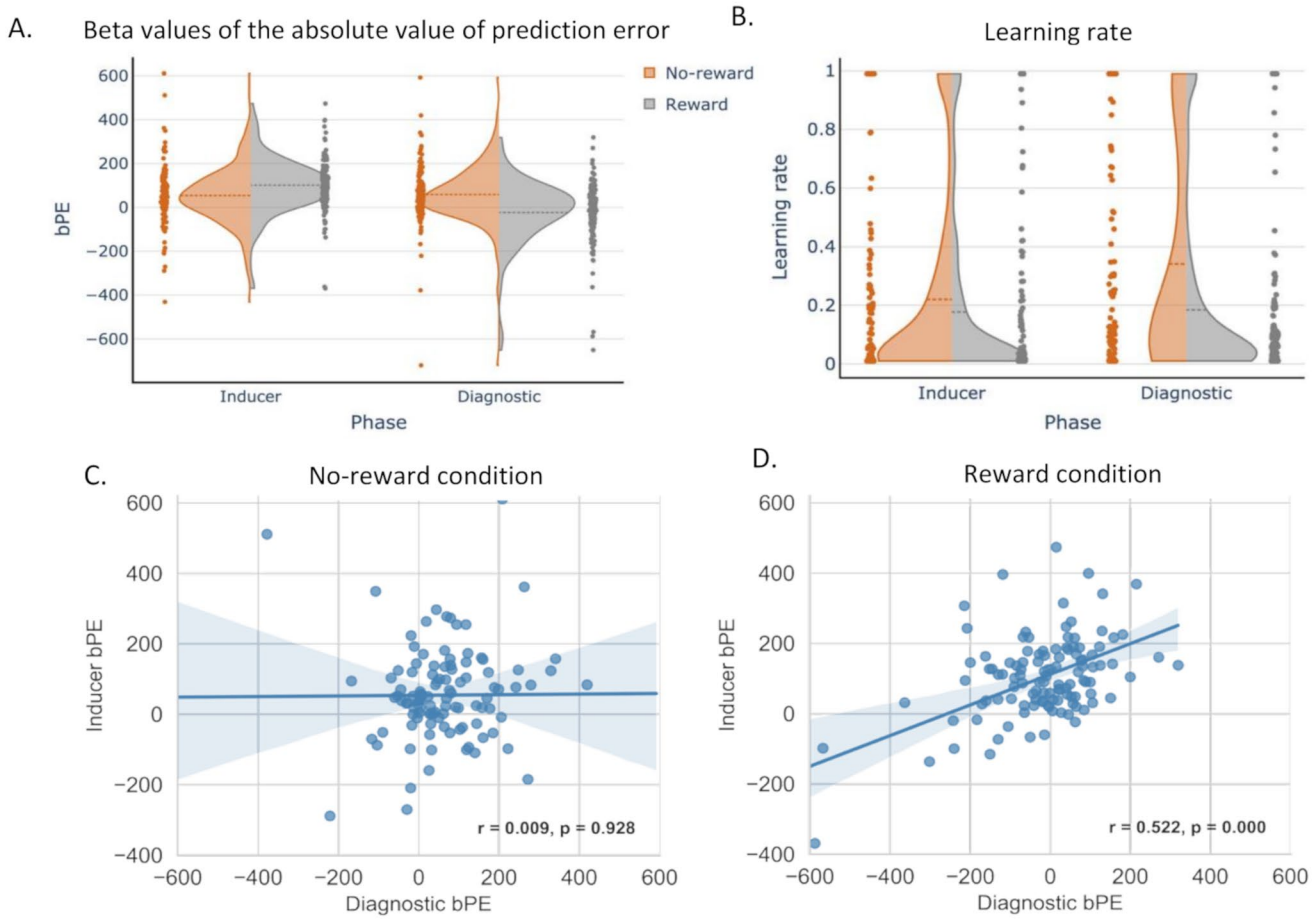

Fig. 3 Model-based RT results. (**A**) Regression weights of the absolute value or prediction error for each condition and each phase. (**B**) Fit learning rates for each condition and phase. (**C**) Correlation between the regression weight of|PE| in the Inducer phase and in the Diagnostic phase for the No-reward condition. (**D**) Correlation between the regression weight of|PE| in the Inducer phase and in the Diagnostic phase for the Reward condition

the non-Gaussian distributions of learning rates a non-parametric test, the Mann-Whitney *U* test (Wilcoxon rank-sum test) was implemented to compare learning rates within the Inducer phase and revealed no significant difference between conditions ($U=6711$, $p=.34$, Common-Language Effect Size (CLES)=0.54; No-reward: $M=0.22$, SE=0.03, Reward: $M=0.17$, SE=0.03, Fig. 3B).

Diagnostic phase. Similar to the Inducer phase, we used β_{PE} as a measure of ISPC learning. In other words, as the learning models adapt to the new ISPC, a strong modulation of PE on RT (i.e., large β_{PE}) indicates that participants learned the (now unbiased) ISPC and used it to influence cognitive control and behavior. Conversely, a weak effect of β_{PE} indicates no learning of ISPC. Thus, by comparing this measure between conditions, we tested whether the presence of reward affects the learning of the new ISPC in the Diagnostic phase. An independent *t*-test showed that the β_{PE} in the No-reward condition ($M=59$, SE=14.0) was larger compared to the Reward condition ($M=$-23, SE=13.72; $t(221)=4.19$, $p<.001$, Fig. 3A). This finding suggests that the Reward condition showed reduced learning of the new (unbiased) ISPC compared to the No-Reward condition. This is consistent with the split-phase RT analysis reported above, in which participants in the Reward condition were shown to carry the outdated ISPC from the Inducer phase throughout the second half of the Diagnostic phase, unlike participants in the No-reward condition.

One alternative explanation to this finding is that the subjects in the Reward condition were in general worse learners of ISPCs. To rule out this possibility, we used the Inducer phase as a baseline and tested the change of β_{PE} from the Inducer to the Diagnostic phase and found that the change in β_{PE} was also significantly larger in the Reward condition ($M=83$, SE=17.95) than the No-reward condition ($M=-0.26$, SE=0.02, $t(213)=$-4.39, $p<.001$, $d=0.58$). This result indicates that subjects in the Reward condition reduced their learning of ISPC compared to those in the No-reward condition.

An additional alternative hypothesis is that the subjects in the Reward condition are more sensitive to the biased ISPC in the Inducer phase (i.e., 90% and 10%) compared to the unbiased ISPC in the Diagnostic phase (i.e., 50%). As β_{PE} scales with the strength of ISPC learning, this hypothesis would predict that subjects showing larger β_{PE} in the Inducer phase will show smaller β_{PE} in the Diagnostic phase. However, opposite to this prediction, we observed a significant positive cross-subject relationship of the PE weights between the two phases in the Reward condition (r=.52, p<.001, Fig. 3C). As a control analysis, we found that when no reward was presented, this relationship was absent in the No-reward condition between phases (r=.009, p=.93, Fig. 3D). This difference in correlations was statistically significant (z=4.19, p<.001). Thus, these findings did not support the alternative explanation.

The Mann-Whitney U test comparing learning rates within the Diagnostic phase revealed a significant difference between conditions (U=7802, p=.001, CLES=0.62; Reward condition: M=0.18, SE=0.03; No-reward condition: M=0.34, SE=0.04, Fig. 3B). The learning rate increased significantly in the No-reward condition from the Inducer phase to the Diagnostic phase $t(214)=2.53$, p=.01, d=0.34, but did not differ significantly between phases for the Reward condition, $t(230)=-0.18$, p=.85, d=0.02. Finally, we found that the change in learning rate between phases was significant between conditions, $t(195)=1.98$, p=.049, d=0.28 (Fig. 3B). The significant increase in the learning rate from the Inducer to the Diagnostic phase for the No-reward study further supports the notion that learning of the ISPC in the Diagnostic phase was faster in the No-reward than the Reward condition. In summary, we observed slower learning of the ISPC in the Diagnostic phase in the Reward than in the No-reward condition, manifested in both PE modulation on RT and learning rates.

Discussion

Extant work probing the influence of reward on cognitive control has examined its contribution to performance in several domains including attention (Botvinick and Braver 2015; Chiew and Braver 2014; Engelmann and Pessoa 2014; Notebaert and Braem 2015; Yee and Braver 2018) and memory encoding (Adcock et al. 2006; Miendlarzewska et al. 2016; Murty et al. 2016; Spaniol et al. 2014; Wittmann et al. 2005). In the present study we aimed to study how cognitive control adapts to multiple external factors, such as reward and statistical demand regularity. Using a between-subject design that manipulated the presence of performance-contingent reward in a variant of the Stroop task, our results first show that, in the Inducer phasev—when no reward was available—participants in both groups learned associations between stimulus features (here, the animals in the pictures) to guide cognitive control. As a result, larger Stroop effects (suggesting less cognitive control) were observed for items that had been MC than MI items, replicating the classic ISPC effect (Bugg et al. 2011; Jacoby et al. 2003). Importantly, in the subsequent Diagnostic phase, where all items were now unbiased, participants in the Reward condition showed slower learning of the new ISPC contingencies than those in the No-reward condition. Specifically, our model-based analyses revealed that participants in the No-reward condition demonstrated a significant increase in β_{PE} from the Inducer phase to the Diagnostic phase, indicating greater sensitivity to environmental changes, whereas participants in the Reward condition did not show a change in β_{PE}, indicating they did not update their control settings. Together, these results provide evidence for the interference of reward incentives on the updating of learned control strategies.

Although the three-way interaction between reward, PC, and congruency was not significant in our factorial analysis, this does not imply that reward did not affect ISPC learning. Instead, our model-based analysis of the PEs revealed that reward influenced ISPC adaptation dynamically rather than as a static interaction. Specifically, we found that reward reduced adaptation to the new (unbiased) ISPC compared to the No-reward condition. This effect was further reflected in learning rates: participants in the No-reward condition exhibited an increase from the Inducer to the Diagnostic phase, whereas no change was observed in the Reward condition.

An alternative explanation is that participants in the Reward condition were simply worse at learning ISPC contingencies. However, we found a significant positive cross-subject correlation between PE weights in the Inducer and Diagnostic phases in the Reward condition, while no such relationship was observed in the No-reward condition. The difference in correlations was statistically significant, ruling out the possibility that individual differences in learning ability drove these effects. Instead, reward appeared to promote stability in learned control settings, leading participants to persist in applying outdated ISPC expectations even when the statistical structure changed.

There are several non-exclusive accounts for why reward hinders the learning of new ISPC contingencies. First, in the Reward condition, participants are simultaneously learning the structure of reward (e.g., the RT threshold to obtain reward, and which items yield high reward, etc.) and the ISPC. Both forms of learning generate prediction error signals in overlapping neural circuits, including the striatum and the midbrain (Daw et al. 2011; Kim et al. 2009; Schönberg et al. 2007; Schultz et al. 1997) and thus may lead to

competition and subsequent masking of the PE caused by the ISPC (Chiu and Egner 2019; Chiu et al. 2017). Second, because learning in the Reward condition involves two factors (reward and ISPC), this may carry an addcasueditional cognitive cost and make the learning of ISPC less effective (Collins 2017). Third, reward may bias attentional resources toward reward-related features at the expense of conflict-driven learning (Niv 2019). It has been shown that attention magnifies PE in perception (Jiang et al. 2013). Similarly, the learning of ISPC, when not attended, may receive only a weak learning signal of PE, leading to reduced learning. Finally, because both reward and ISPC learning occur at the item level, they may compete within working memory, increasing interference and reducing the efficiency of ISPC learning.

Our findings contrast with those of Prével and colleagues (2021), who found that reward reduced the congruency effect when incongruent trials were selectively rewarded but increased the congruency effect when congruent trials were rewarded. Their results suggest that reward can enhance proactive control when it reinforces conflict adaptation (i.e., rewarding incongruent trials), leading to more efficient conflict resolution. Conversely, rewarding congruent trials promoted automatic responses, increasing susceptibility to conflict when incongruent trials appeared.

In contrast, in our study, reward increased the congruency effect. This pattern is more similar to Prével and colleague's (2021) "congruent reward" condition, where reward reinforced reliance on more automated word-reading responses rather than encouraging conflict resolution. However, unlike this study, our reward manipulation did not explicitly reinforce specific trial types—instead, reward was probabilistic and based on overall task performance. This suggests that reward in our study may have stabilized existing control settings rather than promoting flexible adaptation to conflict. Instead of facilitating the ability to update control strategies in response to changing task demands, reward may have reinforced a more rigid approach to conflict processing, amplifying interference from incongruent trials. This interpretation aligns with our major finding that reward masks the learning of new ISPC contingencies, which is another manifestation of stabilized control settings. Just as reward hindered the ability to flexibly adjust ISPC learning in response to new conflict statistics, it may have similarly reinforced pre-existing conflict processing tendencies, increasing rather than decreasing the congruency effect. This provides further evidence that reward can sometimes encourage stability at the cost of flexibility in cognitive control (Chiu and Egner 2019; Chiu et al. 2017; Hefer & Dreisbach., 2017, 2020).

Our findings are also in line with prior empirical work documenting this tradeoff, particularly in tasks that pit cognitive stability against the need for flexible adaptation. Notably, Hefer and Dreisbach (2017, 2020) demonstrated this tradeoff in the AX Continuous Performance Task (AX-CPT). They showed that performance-contingent reward enhanced cue maintenance—reflecting greater cognitive stability—but impaired adaptation when task contingencies changed. This was especially evident when cue validity was removed, yet participants continued to rely on outdated cue-based strategies when behavior was rewarded. This pattern mirrors our findings: in the diagnostic phase, despite the shift in ISPC structure, participants in the Reward condition continued to apply previously learned control settings. Although Hefer and Dreisbach did not use computational modeling, their results suggest that similar reward-driven reductions in learning rate might underlie these effects. Together, their work provides converging support for the idea that reward can foster stable control at the cost of flexibility in adapting to changing task demands.

While our results may seem to be inconsistent with previous research demonstrating that performance-contingent reward increases proactive cognitive control (Chiew and Braver 2014; Fröber and Dreisbach 2014), we argue that both the previous and current findings reflect a similar strategy under the framework of the expected value of control (EVC) theory (Shenhav et al. 2013). The EVC theory posits that cognitive control is applied to balance reward and cost (e.g., cognitive effort). When no reward is available, ISPC can be learned efficiently since effort is strategically allocated where conflict is expected (e.g., when encountering MI items). However, when reward is contingent on performance, cognitive effort allocation is biased towards maximizing reward rather than updating statistical regularities, allowing for more effort to be applied to both MC and MI items. As a result, it is no longer necessary to distinguish between MC and MI items. This is crucial in the current Reward condition, as reward depends on fast responses. Thus, performance-contingent reward will lead to two compatible consequences corresponding to the previous and the current findings. First, cognitive control can be applied proactively (i.e., prior to the onset of the item). Second, the re-learning of ISPC becomes less valuable, as there is no need to separate MC and MI items. This may explain why, in our Reward condition, participants showed greater stability in their control settings rather than flexibly adapting to the new ISPC contingencies.

The ISPC manipulation is designed to demand reactive cognitive control (i.e., demand cannot be predicted before the item is displayed). This leads to the intriguing possibility that this reward-congruence learning tradeoff can be different in the case of proactive cognitive control. For example, in a context-specific proportion congruency task (Jiang et al. 2020b; King et al. 2012), the expected congruence of the

current trial (i.e., the context) can be proactively cued without revealing the correct response. If this paradigm were combined with the current design, we would predict the opposite pattern of results. Specifically, if the present findings reflect a cost-benefit tradeoff in cognitive control, then proactive cognitive control—which can be deployed before the trial start—_would benefit more from the latest congruence information than reactive cognitive control. Therefore, in a rewarded diagnostic phase, participants could improve performance through proactive control, and we would predict increased learning in this condition. Indeed, Braem et al. (2014) found that reward promoted increased sensitivity to context-specific congruency effects, lending credibility to this hypothesis. Though, in these studies, participants were not afforded the opportunity to re-learn previously established statistical regularities in control demands. Alternatively, reduced learning in the rewarded condition (a replication of the current results) would support the hypothesis that reward and congruence PE are encoded by a shared mechanism, and that the reduced learning is driven by their interference.

Another avenue of future research could further investigate whether different reward contingencies (e.g., selectively rewarding incongruent trials) would lead to a reduction in the congruency effect in our paradigm, as observed by Prével and colleagues (2021). If reward selectively enhances proactive control only when directly tied to incongruent trials, this could help resolve the differences between studies and clarify the conditions under which reward promotes flexibility versus stability in cognitive control.

Importantly, a limitation of the present study is its quasi-experimental design: participants in the Reward condition were recruited through a general online platform accessible to the broader population, whereas those in the No-reward condition were recruited from an online university participant pool. Although our model-based analysis helps mitigate concerns about baseline group differences, we caution that differences in recruitment context may limit the generalizability of our findings. Future studies should replicate these results using fully randomized or matched designs.

More broadly, this study made an initial attempt to gauge the relative strength of different factors affecting cognitive control. Real-world control demands often require adaptation to multiple simultaneous factors (e.g., reward, task structure, effort). Although EVC theory has the potential to explain the adaptation of cognitive control, it relies on the understanding of how benefits and costs of multiple factors interact. The interaction seems to be complex, as the factors may not be treated independently. For example, our result shows that reward masks ISPC learning. They indicate that a comprehensive view of how the metacontrol of cognitive control must consider the interaction of multiple factors and provides new insight into how reward modulates the stability-flexibility tradeoff in cognitive control.

Supplementary Information The online version contains supplementary material available at https://doi.org/10.1007/s11571-025-10307-0.

Acknowledgements J.J. is supported by the National Institute of Mental Health (R01MH131559). W.K. is supported by the Office of Naval Research / Department of Defense (N00014-23-1-2792).

Declarations

Open practices statement Data supporting the findings of this study are available on the Open Science Framework (OSF). The dataset can be accessed via the following link: https://osf.io/nxv7t/.

References

Adcock RA, Thangavel A, Whitfield-Gabrieli S, Knutson B, Gabrieli JD (2006) Reward-motivated learning: mesolimbic activation precedes memory formation. Neuron 50(3):507–517

Balleine BW, Dickinson A (1998) Goal-directed instrumental action: contingency and incentive learning and their cortical substrates. Neuropharmacology 37(4–5):407–419

Botvinick M, Braver T (2015) Motivation and cognitive control: from behavior to neural mechanism. Annu Rev Psychol 66:83–113

Botvinick M, Braver TS, Barch DM, Carter CS, Cohen JD (2001) Conflict monitoring and cognitive control. Psychol Rev 108(3):624–652

Braem S, Hickey C, Duthoo W, Notebaert W (2014) Reward determines the context-sensitivity of cognitive control. J Exp Psychol Hum Percept Perform 40(5):1769

Braem S, Bugg JM, Schmidt JR, Crump MJ, Weissman DH, Notebaert W et al (2019) Measuring adaptive control in conflict tasks. Trends Cogn Sci 23(9):769–783

Bugg JM, Crump MJ (2012) In support of a distinction between voluntary and stimulus-driven control: A review of the literature on proportion congruent effects. Front Psychol 3:367

Bugg JM, Dey A (2018) When stimulus-driven control settings compete: On the dominance of categories as cues for control. *Journal of Experimental Psychology: Human Perception and Performance, 44*(12), 1905

Bugg JM, Jacoby LL, Chanani S (2011) Why it is too early to lose control in accounts of item-Specific proportion congruency effects. J Experimental Psychology-Human Percept Perform 37(3):844–859

Bustos B, Colvett JS, Bugg JM, Kool W (2024) Humans do not avoid reactively implementing cognitive control. J Exp Psychol Hum Percept Perform 50(6):587–604

Chaillou A-C, Giersch A, Hoonakker M, Capa RL, Bonnefond A (2017) Differentiating motivational from affective influence of performance-contingent reward on cognitive control: the wanting component enhances both proactive and reactive control. Biol Psychol 125:146–153

Chiew KS, Braver TS (2014) Dissociable influences of reward motivation and positive emotion on cognitive control. Cogn Affect Behav Neurosci 14(2):509–529

Chiu YC, Egner T (2019) Cortical and subcortical contributions to context-control learning. Neurosci Biobehav Rev 99:33–41

Chiu YC, Jiang J, Egner T (2017) The caudate nucleus mediates learning of Stimulus-Control state associations. J Neurosci 37(4):1028–1038
Collins AG (2017) The cost of structure learning. J Cogn Neurosci 29(10):1646–1655
Crump MJC, Gong ZY, Wmiken B (2006) The context-specific proportion congruent Stroop effect: location as a contextual cue. Psychon Bull Rev 13(2):316–321
Daw ND, Gershman SJ, Seymour B, Dayan P, Dolan RJ (2011) Model-based influences on humans' choices and striatal prediction errors. Neuron 69(6):1204–1215
Devine S, Otto AR (2022) Information about task progress modulates cognitive demand avoidance. Cognition 225:105107
Devine S, Roy M, Beierholm U, Otto AR (2024) Proximity to rewards modulates parameters of effortful control exertion. J Exp Psychol Gen
Egner T (2007) Congruency sequence effects and cognitive control. Cogn Affect Behav Neurosci 7(4):380–390
Egner T (2017) The wiley handbook of cognitive control. Wiley Blackwell, Southern Gate, Chichester, West Sussex, UK
Engelmann JB, Pessoa L (2014) Motivation sharpens exogenous spatial attention
Fröber K, Dreisbach G (2014) The differential influences of positive affect, random reward, and performance-contingent reward on cognitive control. Cogn Affect Behav Neurosci 14(2):530–547
Fröber K, Lerche V (2023) Performance-contingent reward increases the use of congruent distracting information. Atten Percept Psychophys 85:905–929
Gratton G, Coles MG, Donchin E (1992) Optimizing the use of information: strategic control of activation of responses. J Exp Psychol Gen 121(4):480
Hefer C, Dreisbach G (2016) The motivational modulation of proactive control in a modified version of the AX-Continuous performance task: evidence from cue-based and prime-based Preparation. Motivation Sci 2(2):116–134
Hefer C, Dreisbach G (2017) How performance-contingent reward prospect modulates cognitive control: increased cue maintenance at the cost of decreased flexibility. J Exp Psychol Learn Mem Cognit 43(10):1643–1658
Hefer C, Dreisbach G (2020) Prospect of performance-contingent reward distorts the action relevance of predictive context information. J Experimental Psychology: Learn Memory Cognition 46(2):380–399
Jacoby LL, Lindsay DS, Hessels S (2003) Item-specific control of automatic processes: Stroop process dissociations. Psychon Bull Rev 10(3):638–644
Jiang J, Summerfield C, Egner T (2013) Attention sharpens the distinction between expected and unexpected percepts in the visual brain. J Neurosci 33(47):18438–18447
Jiang J, Heller K, Egner T (2014) Bayesian modeling of flexible cognitive control. Neurosci Biobehav Rev
Jiang J, Beck J, Heller K, Egner T (2015) An insula-frontostriatal network mediates flexible cognitive control by adaptively predicting changing control demands. Nat Commun 6:8165
Jiang J, Bramao I, Khazenzon A, Wang SF, Johansson M, Wagner AD (2020a) Temporal dynamics of Memory-guided cognitive control and generalization of control via overlapping associative memories. J Neurosci
Jiang J, Wang S-F, Guo W, Fernandez C, Wagner AD (2020b) Prefrontal reinstatement of contextual task demand is predicted by separable hippocampal patterns. Nat Commun 11(1):1–12
Kim H, Sul JH, Huh N, Lee D, Jung MW (2009) Role of striatum in updating values of chosen actions. J Neurosci 29(47):14701–14712
King JA, Korb FM, Egner T (2012) Priming of control: implicit contextual Cuing of top-down attentional set. J Neurosci 32(24):8192–8200
Kool W, Botvinick M (2018) Mental labour. Nat Hum Behav 2(12):899–908
Kool W, McGuire JT, Rosen ZB, Botvinick MM (2010) Decision making and the avoidance of cognitive demand. J Exp Psychol Gen 139(4):665–682
Logan GD, Zbrodoff NJ (1979) When it helps to be misled: facilitative effects of increasing the frequency of conflicting stimuli in a Stroop-like task. Memory Cognition 7:166–174
Miendlarzewska EA, Bavelier D, Schwartz S (2016) Influence of reward motivation on human declarative memory. Neurosci Biobehavioral Reviews 61:156–176
Miller EK, Cohen JD (2001) An integrative theory of prefrontal cortex function. Annu Rev Neurosci 24:167–202
Monsell S (2003) Task switching. Trends Cogn Sci 7(3):134–140
Muhle-Karbe PS, Jiang J, Egner T (2018) Causal evidence for Learning-Dependent frontal lobe contributions to cognitive control. J Neurosci 38(4):962–973
Murty VP, LaBar KS, Adcock RA (2016) Distinct medial Temporal networks encode surprise during motivation by reward versus punishment. Neurobiol Learn Mem 134:55–64
Niv Y (2019) Learning task-state representations. Nat Neurosci 22(10):1544–1553
Notebaert W, Braem S (2015) Parsing the effects of reward on cognitive control. Motivation and cognitive control. Routledge, pp 117–134
Otto AR, Daw ND (2019) The opportunity cost of time modulates cognitive effort. Neuropsychologia 123:92–105
Prével A, Krebs RM, Kukkonen N, Braem S (2021) Selective reinforcement of conflict processing in the Stroop task. PLoS ONE, 16(7), e0255430
Ritz H, Leng X, Shenhav A (2022) Cognitive control as a multivariate optimization problem. J Cogn Neurosci 34(4):569–591
Schmidt JR (2013) Questioning conflict adaptation: proportion congruent and Gratton effects reconsidered. Psychon Bull Rev 20(4):615–630
Schönberg T, Daw ND, Joel D, O'Doherty JP (2007) Reinforcement learning signals in the human striatum distinguish learners from nonlearners during reward-based decision making. J Neurosci 27(47):12860–12867
Schultz W, Dayan P, Montague PR (1997) A neural substrate of prediction and reward. Science 275(5306):1593–1599
Shenhav A, Botvinick MM, Cohen JD (2013) The expected value of control: an integrative theory of anterior cingulate cortex function. Neuron 79(2):217–240
Shenhav A, Musslick S, Lieder F, Kool W, Griffiths TL, Cohen JD et al (2017) Toward a rational and mechanistic account of mental effort. Annu Rev Neurosci
Spaniol J, Schain C, Bowen HJ (2014) Reward-enhanced memory in younger and older adults. Journals Gerontol Ser B: Psychol Sci Social Sci 69(5):730–740
Spinelli G, Morton JB, Lupker SJ (2022) Both task-irrelevant and task-relevant information trigger reactive conflict adaptation in the item-specific proportion-congruent paradigm. Psychon Bull Rev, 1–13
Stroop JR (1935) Studies of interference in serial verbal reactions. J Exp Psychol 18(6):643
Suh J, Bugg JM (2021) On the automaticity of reactive item-specific control as evidenced by its efficiency under load. J Exp Psychol Hum Percept Perform 47(7):908
Sutton RS, Barto AG (2018) *Reinforcement learning: an introduction* (Second edition. ed.). Cambridge, Massachusetts: The MIT Press
Ullsperger M, Bylsma LM, Botvinick MM (2005) The conflict adaptation effect: it's not just priming. Cogn Affect Behav Neurosci 5(4):467–472
Vallat R (2018) Pingouin: statistics in Python. J Open Source Softw 3(31):1026

Westbrook A, Kester D, Braver TS (2013) What is the subjective cost of cognitive effort? Load, trait, and aging effects revealed by economic preference. PLoS ONE, 8(7), e68210

Wittmann BC, Schott BH, Guderian S, Frey JU, Heinze H-J, Düzel E (2005) Reward-related FMRI activation of dopaminergic midbrain is associated with enhanced hippocampus-dependent long-term memory formation. Neuron 45(3):459–467

Yee DM, Braver TS (2018) Interactions of motivation and cognitive control. Curr Opin Behav Sci 19:83–90