



Control by association: Transfer of implicitly primed attentional states across linked stimuli

Christina Bejjani^{1,2} · Ziwei Zhang¹ · Tobias Egner^{1,2}

Published online: 15 February 2018
© Psychonomic Society, Inc. 2018

Abstract

Although cognitive control has traditionally been viewed in opposition to associative learning, recent studies show that people can learn to link particular stimuli with specific cognitive control states (e.g., high attentional selectivity). Here, we tested whether such learned stimulus-control associations can transfer across paired-associates. In the Stimulus-Stimulus (S-S) Association phase, specific face or house images repeatedly preceded the presentation of particular scene stimuli, creating paired face/house-scene associates in memory. The Stimulus-Control (S-C) Association phase then associated these scenes with different attentional control states by probabilistically biasing specific scenes to mostly precede either congruent or incongruent trials in a Stroop task. Finally, in the Stimulus-Control Transfer (S-CT) phase, the faces and houses from the S-S phase preceded Stroop trials but were not predictive of congruency, testing whether stimulus-control associations would transfer from scenes to their associated face/house stimuli. In Experiments 1 and 3, we found that learned implicit stimulus-control associations could transfer across closely linked cues, and in Experiment 2, we showed that this transfer depended on the memory associations formed in the S-S phase. While this form of transfer learning has previously been demonstrated for stimulus-reward associations, the present study provides the first evidence for the associative transfer of stimulus-control associations across arbitrarily linked stimuli. This work demonstrates how people can learn to implicitly adapt their processing strategies in a flexible context-dependent manner and establishes a novel learning mechanism supporting the generalization of cognitive control.

Keywords Cognitive control · Control learning · Attention · Memory

Introduction

Cognitive control describes the strategic use of a current context to guide our responses in line with internal goals. Traditionally, control is viewed as an opposing force to bottom-up associational processes (i.e., learning and memory)

that drive automatic, overlearned responses (Norman & Shallice, 1986). For instance, in the color-word Stroop task (Stroop, 1935), participants must respond to the ink color of color-words while suppressing habitual word-reading, and are slower and less accurate when the meaning of the word and its ink color are incongruent (e.g., RED in blue) than when they are congruent (e.g., RED in red). This Stroop effect is thought to reflect the behavioral cost of exerting control to resolve the clash between the habitual reading response and the contextually appropriate instructed ink-naming response (Botvinick, Braver, Barch, Carter, & Cohen, 2001). However, such juxtaposition between associative and controlled processing ignores the fact that learning and attentional control need to work hand-in-hand so that contexts can be linked with appropriate control settings (Egner, 2014).

This type of context-control learning has been demonstrated in a number of studies where implicit probabilistic cues (e.g., stimulus location, color, trial type history, or sensory modality) have been found to facilitate the retrieval of context-appropriate attentional control states (e.g., high

Electronic supplementary material The online version of this article (<https://doi.org/10.3758/s13423-018-1445-6>) contains supplementary material, which is available to authorized users.

✉ Christina Bejjani
christina.bejjani@duke.edu

¹ Department of Psychology and Neuroscience, Duke University, Durham, NC 27708, USA

² Center for Cognitive Neuroscience, Duke University, LSRC Box 90999, Durham, NC 27708, USA

attentional focus) (for reviews, see Bugg, 2017; Bugg & Crump, 2012). For instance, Stroop stimuli presented at a specific location (e.g., top-right of screen) that is mostly associated with higher control-demand (frequent incongruent trials) are processed more efficiently (producing smaller Stroop effects) relative to those presented in a location that is mostly associated with lower control-demand (frequent congruent trials). Contextual cues can thus guide strategic adjustment to control demands. Such stimulus-control learning is highly adaptive, combining the speed of automatic processing with the flexibility of cognitive control (Egner, 2014).

Yet greater flexibility could be achieved if learned control states were transferred across associated stimuli or contexts. This type of transfer has been shown to occur for stimulus-reward associations: when stimulus A predicts stimulus B, and stimulus B is subsequently paired with reward, participants transfer B's reward value to A (Wimmer & Shohamy, 2012). Previously, context-specific (e.g., location-driven) control settings have been shown to extend to unbiased and novel stimuli (e.g., Crump & Milliken, 2009; King, Korb, & Egner, 2012; Surrey, Dreisbach, & Fischer, 2017; Weidler & Bugg, 2016) as well as to members within particular social categories (Cañadas, Rodríguez-Bailón, Milliken, & Lupiáñez, 2013), but whether control-state associations can transfer across closely linked cues remains unknown. Transferring attentional settings across associated stimuli would provide a valuable means of generalizing learned control states and thus their associated cognitive strategies. Here, in three experiments, we tested stimulus-to-stimulus transfer learning of cognitive control.

Experiment 1

Our goal was to test whether learned stimulus-control state associations would transfer across paired-associates. Our task consisted of a stimulus-stimulus (S-S) association phase, a stimulus-control (S-C) association phase, and a stimulus-control transfer (S-CT) test phase (cf. Wimmer & Shohamy, 2012). In the S-S phase, face or house (S1) images repeatedly preceded the presentation of scene (S2) images, in order to foster S1-S2 paired-associates in memory. Next, in the S-C phase, S2 images preceded color-word stimuli in a Stroop task. Crucially, S2 images served as implicit probabilistic cues: half were predictive of congruent trials (low control-demand) and the other half of incongruent trials (high control-demand). Finally, in the S-CT phase, S1 images, instead of S2 images, preceded the Stroop stimuli, but did not have any predictive relationship with congruency. This allowed us to test whether any S2 control-demand associations acquired during the S-C phase would transfer to their S1 paired-associates. Control transfer learning would be evident if S1 images associated with S2 high control-demand images produced a

smaller congruency effect in the transfer phase compared to S1 images associated with S2 low control-demand images.

Method

Participants

Our target sample size was 75 participants based on effect sizes ($\eta_p^2 = 0.11$, $\eta_p^2 = 0.17$) from a recent study on context-control learning (Surrey et al., 2017). With a high power of 0.9 and a Type I error level of 0.05, we required between 54 and 87 participants, and we anticipated an attrition rate of ~10%. We therefore aimed to run ~80 participants.

Eighty-three Amazon Mechanical Turk (MTurk) workers consented to participate for a US\$7.50 fee. Seven participants were excluded because of poor accuracy (i.e., less than 70% in at least one task phase), resulting in a final sample size of 76 (mean age = 34.01 years, SD = 8.88; 42 female, 33 male, one no reply). Prior to accepting the HIT, participants were told the amount of time that each task phase would take and that they would have to categorize images and make judgments about color-words.

Experimental procedure

The experiment consisted of consecutive S-S, S-C, and S-CT phases (Fig. 1). On each trial in the S-S phase (Fig. 1b), participants saw a face or house (S1) image for 750 ms, a fixation cross for 500 ms, and a scene (S2) image for 750 ms, followed by performance feedback (correct/incorrect) for 1,000 ms and a blank ITI screen for 1,500 ms. The participants' instructed goal was to categorize S2 images by pressing scene-specific keyboard response buttons (a, s, k, and l on a QWERTY keyboard, mapped onto left middle, left index, right index, and right middle fingers, respectively). However, each S2 image was preceded by a specific S1 image with a 0.8 probability and for the remaining 20% of trials, was preceded equally often by each of the other three invalid S1 images. Participants could thus optimize their performance by learning and exploiting these S1-S2 paired-associates.

To establish an association between S2 images and attentional control states, participants then underwent the S-C phase (Fig. 1c). Here, participants performed the color-word Stroop task (Stroop, 1935), using four color-words (BLUE, GREEN, RED, YELLOW) printed either in congruent (e.g., RED in red) or incongruent (e.g., RED in blue) colors. Participants responded to the print color via color-specific response buttons (z, x, n, and m on a QWERTY keyboard, mapped onto left middle, left index, right index, and right middle finger, respectively). S2 images were shown for 1,000 ms, followed by a fixation cross for 1,000 ms, a Stroop color-word for 1,000 ms, and an ITI of 2,000 ms.

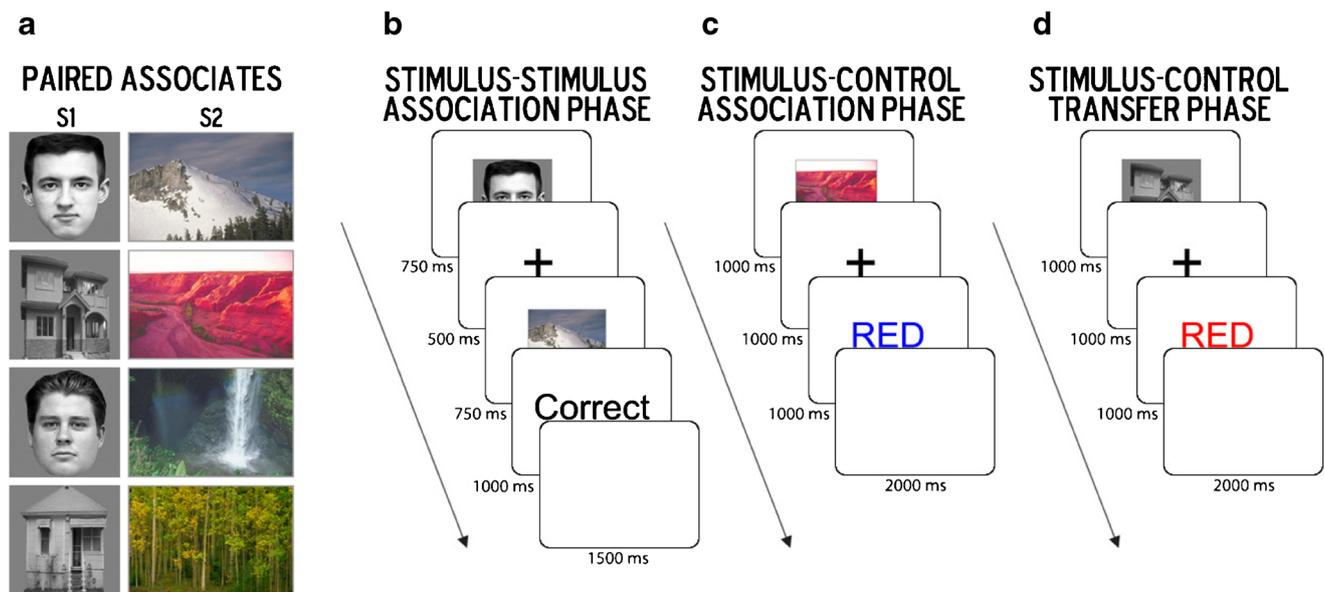


Fig. 1 Summary of task procedure. (a) Stimuli consisted of two grayscale male face images from Kanade, Cohn, and Tian (2000), two grayscale house images from real estate websites, and four scene images (mountain, canyon, waterfall, forest). Face and house images were used for their known neural selectivity, in anticipation of running an fMRI study at a later date. (b) In the Stimulus-Stimulus (S-S) association phase,

a face or house (S1) repeatedly predicted a particular scene (S2) image to form paired associates in memory. (c) In the Stimulus-Control (S-C) association phase, S2 images predicted stimulus congruency in a Stroop task to create implicit "control-demand cues." (d) In the Stimulus-Control Transfer (S-CT) phase, S1 "transfer probes" likewise preceded the onset of Stroop trials but were not predictive of congruency

While the overall frequency of congruent/incongruent trials was equal, two S2 images mostly preceded incongruent trials, while the other two mostly preceded congruent trials (both at 0.8 probability). Our manipulation thus aimed to create two types of S2 "control-demand cues": one associated with a low state of attentional control (congruent trial cues) and another associated with a high state of attentional control (incongruent trial cues).

Crucially, in the S-CT phase (Fig. 1d), we then tested whether the attentional control states linked with S2 control-demand cues would transfer to their S1 paired-associate. Participants again performed the color-word Stroop task, but with S1 images preceding the color-word trials. Importantly, unlike S2 control-demand cues, S1 images were not predictive of congruency (0.5 probability). Any performance differences observed as a function of these S1 "transfer probes" could thus be attributed to the transfer of stimulus-control associations from S2 images to their respective S1 paired-associate. The S-CT phase followed the same trial timing as the S-C phase.

For all task phases, if participants did not respond before the scene (S-S) or color-word (S-C and S-CT) disappeared from the screen, a feedback time-out was provided for 1,000 ms to encourage quicker responses. This occurred on a total of 5.98 % of S-S phase trials, 3.06 % of S-C phase trials, and 3.07 % of S-CT phase trials. A coding error resulted in the data exclusion of the last three to four S-S phase trials (0.47 % total) for 27 participants.

All S1 and S2 images were presented in the center of the screen and with equal frequency in every task phase (S-S

phase: 240 total trials across four runs; S-C: 320 trials across four runs; S-CT: 160 trials across two runs), and all trial types in each phase were presented in random order.

To counterbalance the assignment of paired-associates, we ran four task versions. In two versions, the two faces and two houses were associated with the scenes that represented low/high control-demand, respectively, while in the other two versions, one face and one house each were associated with low/high control-demand, respectively. We did not counterbalance for response button assignment, since any random bias in response fluency to a particular stimulus would be orthogonal to our effects of interest.

Post-test questionnaire

After the main experiment, participants filled out a questionnaire that assessed their memory of the S-S phase paired-associates and explicit understanding of the task structure (see [Supplementary Material](#)). Note that participants did not know in advance that they would be asked about the task structure or to recall the S-S phase associates.

Data analysis

Analyses were carried out on accuracy (proportion correct) and reaction time (RT) data. RT was analyzed for correct trials that did not entail a direct stimulus repetition from the previous trial (S-S: 0.39 % of trials; S-C: 0.11 %; S-CT: 0.13 %)

and that were not excessively fast (< 200 ms) or slow (feedback time-out S-S: > 750 ms; S-C/S-CT: > 1,000 ms).

For the S-S phase, we compared trials in which a S1 image preceded its S2 paired-associate (valid) to those in which it did not (invalid) using a paired-samples *t*-test. For the S-C phase, a $2 \times 2 \times 4$ repeated-measures ANOVA was run, with S2 control-demand cue (low/high), congruency (congruent/incongruent), and run (1–4) as within-subjects factors. The S-CT phase data were analyzed similarly, with S1 transfer probes replacing the corresponding S2 control-demand cues and with only two runs. The run factor was of particular interest for the S-CT phase, since the S1 transfer probes were not predictive of congruency and any transfer effect was therefore expected to be short-lived. All data were Greenhouse-Geisser corrected where appropriate.

All materials are online at <https://github.com/christinabejjani/controltransfer/expt1>.

Results

Stimulus-Stimulus (S-S) association phase

As anticipated, participants formed S1-S2 associations: validly cued S2 images were categorized more quickly (Fig. 2a; $t(75) = 4.01, p < 0.001$, Cohen's $d = 0.42$, CL effect size = 68 %) and accurately ($t(75) = -2.21, p = 0.030$, Cohen's $d = 0.27$, CL effect size = 60 %) than invalidly cued ones. See Table S1 for full RT and accuracy data.

Stimulus-Control (S-C) association phase

Here, S2 images were used as "control-demand cues": half mostly predicted congruent Stroop trials (low control-demand), while the other half mostly predicted incongruent Stroop trials (high control-demand). As anticipated, participants responded more slowly and less accurately on incongruent than congruent trials ($F(1,75) = 402.82, p < 0.001, \eta_p^2 = 0.84$; $F(1,75) = 104.66, p < 0.001, \eta_p^2 = 0.58$). Surprisingly, at the group level, we did not observe any evidence that participants used the S2 control-demand cues to modulate Stroop congruency (Fig. 2b; main effect of S2 control-demand cue: $F_s < 1$; cue \times congruency $F_s < 1$; run \times cue \times congruency RT: $F < 1$, accuracy: $F(3,225) = 1.56, p = 0.199, \eta_p^2 = 0.02$). However, there was a large amount of individual variability in S-C cueing effects (see S-CT phase section).

Because the S-S phase could have promoted the categorization of scenes associated with faces versus those associated with houses and thus impeded learning in the S-C phase, we also tested whether the specific stimulus pairings across the different counterbalancing versions of the task impacted S-C learning. There was no effect of task version on the S-C cueing

effects (high vs. low control-demand congruency effect) for RT ($t(74) = 0.32, p = 0.754$) or accuracy ($t(74) = 0.20, p = 0.840$).

Stimulus-Control Transfer (S-CT) phase

In line with the assumption that S-C associations would transfer from S2 control-demand cues to their paired-associate S1 transfer probes, we observed a run-sensitive transfer RT effect: the mean congruency effect was reduced for high compared to low control-demand S1 transfer probes in run 1 (Fig. 2c; $F(1,75) = 4.89, p = 0.030, \eta_p^2 = 0.06$) but not in run 2 (run 2 cue \times congruency: $F(1,75) = 1.95, p = 0.167, \eta_p^2 = 0.03$; run \times cue \times congruency: $F(1,75) = 5.76, p = 0.019, \eta_p^2 = 0.07$). We observed no transfer accuracy effects (run 1 cue \times congruency: $F < 1$; run 2 cue \times congruency: $F(1,75) = 2.03, p = 0.159, \eta_p^2 = 0.03$; run \times cue \times congruency: $F(1,75) = 1.51, p = 0.223, \eta_p^2 = 0.02$).

Given the absence of a group level cueing effect in the S-C phase, a control transfer RT effect implies that a subset of individuals acquired S2 control-demand associations in the S-C phase. We therefore examined whether individual differences in S-C learning could explain the presence of a RT transfer effect. Indeed, as the S-C RT cueing effect (high vs. low control-demand congruency effect) increased, so did the S-CT RT effect ($r = 0.23, p = 0.045$). We divided the S-C and S-CT effects for each participant by the mean RTs in each phase to control for overall differences in RT, but the S-C/S-CT RT relationship was unaltered ($r = 0.24, p = 0.035$). Similarly, when we categorized participants as S-C non-learners (i.e., negative RT cueing effect) and S-C learners (i.e., positive RT cueing effect), we found that S-C non-learners did not show a transfer effect (runs 1 and 2 cue \times congruency: $F_s < 1.1$; run \times cue \times congruency: $F < 1$), while S-C learners did (run 1 cue \times congruency: $F(1,37) = 6.24, p = 0.017, \eta_p^2 = 0.14$; run 2 cue \times congruency: $F < 1$; run \times cue \times congruency: $F(1,37) = 5.28, p = 0.027, \eta_p^2 = 0.13$). This suggests that the attentional states learned in the S-C phase drove the observed transfer effect through the S1-S2 association.

Memory of S1-S2 associative pairs

Participants performed better than chance ($M = 2.45, 95\% \text{ CI } 2.15\text{--}2.75$) when matching a S1 image with one of two S2 options ($t(75) = 2.93, p = 0.004$, Cohen's $d = 0.34$), and memory performance was correlated with the S-S validity effects for RT ($r = 0.46, p < 0.001$) and accuracy ($r = -0.38, p = 0.001$). Thus, participants who learned the initial S1-S2 associations had greater subsequent memory of these pairs.

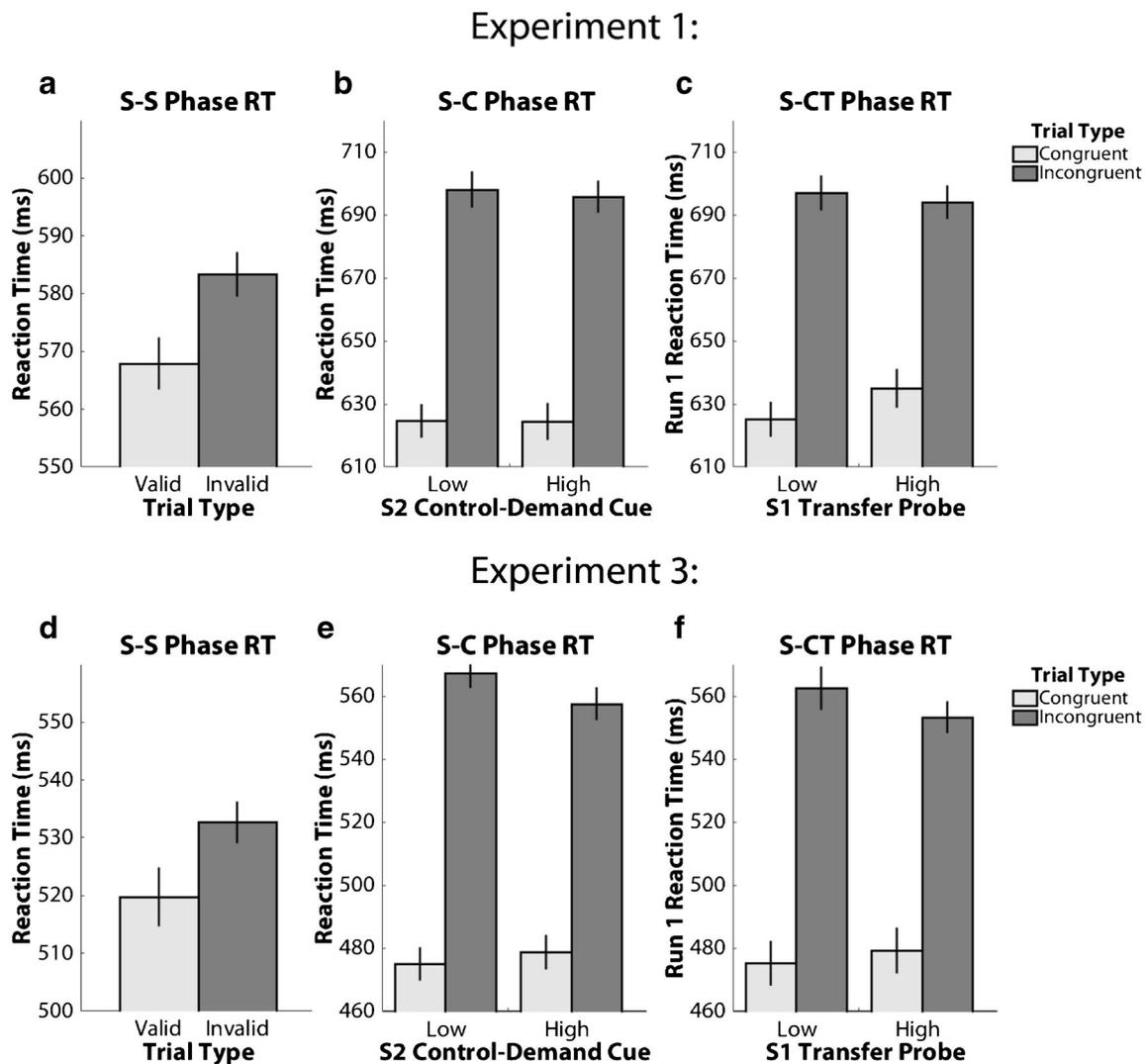


Fig. 2 Experiments 1 and 3 results. (a and d) S-S phase mean RT (\pm SEM) data are displayed as a function of S1-S2 cue validity. (b and e) S-C phase Stroop task performance: Mean RT (\pm SEM) data are plotted as

a function of congruency and S2 control-demand cue. (c and f) S-CT phase: run 1 mean RT (\pm SEM) data are shown as a function of congruency and S1 transfer probe

Explicit understanding of task structure

Given that participants remembered the S1-S2 pairs, we assessed whether they had explicit knowledge of the S-S task structure through a series of post-test questions (see [Supplementary Material](#)). Only four participants reported noticing the underlying task structure and answered all follow-up questions correctly, indicating that most participants had limited explicit awareness of the S-S phase structure.

We assessed explicit awareness of the underlying S-C and S-CT task structures by asking participants whether they noticed any systematic variation in the image/Stroop trial sequences. Only three participants exhibited any explicit understanding of the S-C phase structure. Removing these subjects from the main analyses did not alter the results. Participants also rated the extent to which image categories predicted the

frequency of Stroop trial congruency and then matched scenes, faces, and houses to their most likely trial type. Participants were unable to differentiate between the actual predictor category and lures, and match the S2 cues to their respective congruencies. This suggests that most participants remained unaware of the task structure and learned these associations implicitly.

Experiment 2

We interpret the results of Experiment 1 as a demonstration of an implicit stimulus-control transfer effect, where control state associations are passed between previously established paired-associates. Alternatively, however, the S-CT phase effect could have arisen due to “learned predictiveness” (cf. Le Pelley & McLaren, 2003),

whereby participants learned that (scene) image cues predicted Stroop trial congruency in the S-C phase and subsequently assumed that the same predictive structure would hold true for the S-CT phase, where the face/house images act as cues for Stroop trials in a similar task structure. If participants made this assumption about the predictiveness of face/house images in the S-CT phase, this account would also predict a run-sensitive transfer effect, because participants would soon realize that the learned structure was not a helpful performance predictor. Moreover, such an account tests the foundational assumption of this study: that learned control states are transferred across the initial S-S associations.

In Experiment 2, we tested this rival account by scrambling the S1-S2 associations, such that no paired-associates could be formed. Here, the control-transfer hypothesis would predict that no differential congruency effects should be observed in the S-CT phase, whereas the learned predictiveness account would still anticipate such effects.

Method

Participants

Seventy-nine MTurk workers consented to participate for a US\$7.50 fee. Six participants were excluded because of poor accuracy (< 70 %), resulting in a final sample size of 73 (mean age = 33.26 years, SD = 9.07; 44 female, 29 male). HIT instructions remained the same as in Experiment 1.

Experimental procedure

Only the S-S phase differed from the Experiment 1 procedure. Here, S1 images preceded each S2 image at an equal rate. Across all participants, a feedback time-out occurred on 6.43 % of S-S phase trials, 3.97 % of S-C phase trials, and 3.46 % of S-CT phase trials.

Since the S-S phase no longer included paired-associates, we ran a 4×4 repeated-measures ANOVA with S1 and S2 images as factors to determine the effect of S1 cues on S2 categorization. The S-C phase data were analyzed as in Experiment 1. For S-CT data, to test for the presence of a transfer effect, we performed a $4 \times 2 \times 2$ repeated-measures ANOVA with S1 images, congruency, and run as factors. Because pairs of S1 cues were analyzed together in Experiment 1, we also ran paired t-test comparisons of congruency effects between all pairs of S1 images.

All materials are online at <https://github.com/christinabejjani/controltransfer/expt2>.

Results

Stimulus-Stimulus (S-S) association phase

Now that each S1 image preceded each S2 image an equal number of times, S1 images did not impact S2 categorization (main effect of S1 image: $F_s < 1$; S1 \times S2 image: $F_s < 1.1$). There was an unexpected main effect of S2 image (RT: $F(3,216) = 20.89, p < 0.001, \eta_p^2 = 0.23$; accuracy: $F(3,216) = 17.20, p < 0.001, \eta_p^2 = 0.19$), such that participants were faster to categorize the forest and more accurate at categorizing the mountain than other scenes. This is likely due to low-level image features (e.g., the forest being the brightest stimulus) and should not impact S-C and S-CT Stroop performance. See Table S2 for full RT and accuracy data.

Stimulus-Control (S-C) association phase

The S-C phase replicated that of Experiment 1: participants were faster ($F(1,71) = 339.63, p < 0.001, \eta_p^2 = 0.83$) and more accurate ($F(1,72) = 84.13, p < 0.001, \eta_p^2 = 0.54$) at responding to congruent stimuli. Their responses were also not impacted by S2 control-demand cues ($F_s < 1.4$), and there was no interaction between run, control-demand cue, and congruency for RT ($F(3,213) = 1.79, p = 0.151, \eta_p^2 = 0.03$). However, this three-way interaction was significant for accuracy ($F(3,216) = 2.88, p = 0.043, \eta_p^2 = 0.04$). Post-hoc analysis suggests that this occurred due to a linear trend ($F(1,72) = 5.73, p = 0.019, \eta_p^2 = 0.07$), driven by the decrease in congruency effects for the low control-demand cue over time, while congruency effects for the high control-demand cue remained steady over time.

Stimulus-Control Transfer (S-CT) phase

If the Experiment 1 results were due to transfer learning, no transfer should be observed in Experiment 2. However, if the Experiment 1 results were caused by learned predictiveness, we would expect differential congruency effects in the Experiment 2 S-CT phase. Consistent with the control transfer hypothesis, there was no interaction between S1 image and congruency, regardless of whether we focused only on run 1 RT/accuracy ($F_s < 1$), or included the run factor in the S1 \times congruency ANOVA ($F_s < 1$). Similarly, when we compared congruency effects between all pairs of S1 stimuli, we observed no effects (all t -stats < 1.05).

Experiment 3

Taken together, the results of Experiments 1 and 2 suggest that control transfer learning is possible. However, we observed no mean group effect of control-learning in the S-C phase, and

the transfer effect was modest in size. Moreover, one possible alternative account for the transfer effect in Experiment 1 is that it might reflect the transfer of simple S-R contingencies rather than of control settings (see [Supplementary Material](#)). We therefore designed a third experiment with the following goals: First, we aimed to conceptually replicate the control transfer learning effect of Experiment 1. Second, we sought to enhance the S-C phase learning effect, both to foster a group-level effect and as a potential means to enhance transfer. To this end, we aimed to amplify Stroop conflict by increasing the time between the presentation of a Stroop color-word and its printed color (cf. Appelbaum, Meyerhoff, and Woldorff, 2009), and we sought to strengthen the S-C Phase associations between scene images and congruency by showing the S2 images for the entire trial duration (with the Stroop color-word shown underneath the scene image). Third, we aimed to rule out the possibility of a contingency learning confound. For this purpose, we manipulated the Stroop stimulus set in the S-C Phase such that a subset of “biased” color words created the cue-congruency associations, while another subset of color-words served as unbiased test items to index stimulus-control learning in the absence of stimulus-response learning confounds (Crump & Milliken, 2009). Finally, since Experiments 1 and 2 were conducted online, we also wanted to ensure the transfer effect would be observable under standard laboratory conditions. Thus, Experiment 3 was run in-person.

Method

Participants

Given our assumption that the modifications to the S-C Phase in Experiment 3 would increase the effect sizes of control-learning and transfer effects, we reduced our target sample to 45 participants. Based on the attrition rate of the previous experiments, we aimed to recruit ~50 participants.

Fifty-two undergraduate students consented to participate in a 90-min slot for course credit and were not told about the number of task phases or their length prior to starting the study. Eight participants were excluded because of poor accuracy (< 70 %), resulting in a final sample size of 44 (mean age = 18.89 years, SD = 1.30; 34 female, 10 male).

Experimental procedure

In comparison to Experiment 1, we altered the S-C and S-CT phases. To ensure that participants would use S2 images as control-demand cues, the scenes were presented for the entire 3-s trial duration 100 pixels above the center of the screen, while the fixation cross and the Stroop stimuli that followed were presented 100 pixels below the center of the screen. The

Stroop stimulus timing differed with one regard to that in Experiment 1: a color-word was presented for 200 ms in black before its response-eligible printed color was presented for the remaining 800 ms.

To control for the formation of stimulus-response associations between scene stimuli and specific Stroop items (which might in theory transfer to the associated S1 images), we split Stroop stimuli into two sub-sets, frequency-biased and frequency-unbiased items (cf. Crump & Milliken, 2009). Specifically, the words red and green (frequency-biased stimuli) were always incongruent following high control-demand cues and always congruent following low control-demand cues, and were only printed in either red or green. By contrast, the words blue and yellow (frequency-unbiased stimuli) were incongruent or congruent half of the time, not predicted by the cue, and were only printed in either blue or yellow. This ensured that control-demand cues mostly preceded their respective trial types at a 0.75 probability while also producing “unbiased” probe items (the blue/yellow stimuli) that indexed an effect of learned control in the absence of the confound of learned stimulus-response associations. Thus, our key question was whether these unbiased Stroop items would show (a) an S-C learning effect in the S-C Phase, and, most importantly, (b) a control transfer effect in the S-CT Phase.

Across all participants, a feedback time-out occurred on 6.40 % of S-S phase trials, 4.22 % of S-C phase trials, and 3.92 % of S-CT phase trials. The S-C phase also now consisted of 256 total trials. The S-CT phase followed the structure of the S-C phase. Analyses were performed on the frequency-unbiased and frequency-biased stimuli separately.

All materials are online at <https://github.com/christinabejjani/controltransfer/exp3>.

Results

Stimulus-Stimulus (S-S) association phase

Consistent with Experiment 1, participants formed S1-S2 associations: validly cued S2 images were categorized more quickly (Fig. 2d; $t(43) = 3.05$, $p = 0.004$, Cohen's $d = 0.44$, CL effect size = 68 %) and slightly more accurately ($t(43) = 1.85$, $p = 0.071$, Cohen's $d = 0.23$, CL effect size = 61 %) than invalidly cued ones. See Table S3 for full RT and accuracy data.

Stimulus-Control (S-C) association phase

Now that we increased Stroop conflict and showed the S2 control-demand cues for the entire duration of each trial, participants used these cues to modulate Stroop congruency for the frequency-unbiased stimuli (Fig. 2e; cue × congruency:

$F(1,43) = 10.55, p = 0.002, \eta_p^2 = 0.20$; main effect of cue: $F < 1.6$ and congruency: $F(1,43) = 361.15, p < 0.001, \eta_p^2 = 0.89$; run \times cue \times congruency: $F < 1$). The difference in congruency for frequency-biased stimuli was smaller than congruency differences for frequency-unbiased stimuli following both the high-demand cues ($t(43) = 3.34, p = 0.002$, Cohen's $d = 0.53$, CL effect size = 69 %) and low-demand cues ($t(43) = 6.03, p < 0.001$, Cohen's $d = 1.01$, CL effect size = 82 %). This suggests that participants learned stimulus-response associations, but that these associations did not drive our conditioned control-demand effect.

As with Experiment 1, we did not observe these effects for accuracy (cue \times congruency: $F(1,43) = 2.69, p = 0.108, \eta_p^2 = 0.059$; main effect of cue: $F < 1$, congruency: $F(1, 43) = 21.16, p < 0.001, \eta_p^2 = 0.330$; run \times cue \times congruency: $F < 1$; frequency-biased vs. high-demand: $t(43) = 0.90, p = 0.374$, Cohen's $d = 0.17$, CL effect size = 55 %; frequency-biased vs. low-demand: $t(43) = 0.69, p = 0.496$, Cohen's $d = 0.14$, CL effect size = 54 %).

Stimulus-Control Transfer (S-CT) phase

Experiment 3 successfully replicated the run-sensitive control-transfer effect of Experiment 1: the mean congruency effect was reduced for high compared to low control-demand S1 transfer probes in run 1 (Fig. 2f; cue \times congruency: $F(1,43) = 4.59, p = 0.038, \eta_p^2 = 0.10$) but not in run 2 (cue \times congruency: $F < 1.1$; run \times cue \times congruency: $F(1,43) = 4.76, p = 0.035, \eta_p^2 = 0.10$) for frequency-unbiased stimuli. This effect was not observed for frequency-biased stimuli (all F s < 1). Thus, while stimulus-control learning effects were transferred, stimulus-response learning effects were not. We observed no transfer accuracy effects (all frequency-unbiased F s < 1 ; frequency-biased: run \times cue \times congruency: $F < 1$; run 1 cue \times congruency: $F < 1$; run 2 cue \times congruency: $F(1,43) = 1.49, p = 0.229, \eta_p^2 = 0.033$).

As with Experiment 1, we found that across individuals, the S-C learning effect (high – low control-demand cue congruency difference, for frequency-unbiased stimuli) was correlated with the run-sensitive, frequency-unbiased transfer effect ($r = 0.34, p = 0.024$), even when we corrected for mean differences in RT within each participant ($r = 0.34, p = 0.023$). These relationships did not exist for frequency-biased stimuli ($r = 0.13, p = 0.401$; corrected: $r = 0.15, p = 0.317$).

Confirming the individual differences in transfer learning we observed in Experiment 1, we also found that RT transfer was stronger in S-C learners (run \times cue \times congruency: $F(1,30) = 7.58, p = 0.010, \eta_p^2 = 0.20$; run 1 cue \times congruency: $F(1,30) = 21.14, p < 0.001, \eta_p^2 = 0.41$; run 2 cue \times congruency: $F < 1$) than S-C non-learners (run \times cue \times congruency: $F < 1$; run 1 cue \times congruency: $F < 1$; run 2 cue \times congruency: $F(1,12) = 3.37, p = 0.091, \eta_p^2 = 0.22$).

Memory of S1-S2 associative pairs

Replicating Experiment 1, participants performed better than chance ($M = 2.64$, 95 % CI 2.24–3.03) when matching S1 and S2 images ($t(43) = 3.13, p = 0.003$, Cohen's $d = 0.47$). Memory performance was also correlated with the S-S validity effects for RT ($r = 0.45, p = 0.002$) but not accuracy ($r = -0.15, p = 0.339$).

Explicit understanding of task structure

Replicating Experiment 1, most participants indicated limited explicit awareness of the underlying S-S, S-C, and S-CT phase structures (see [Supplementary Material](#)). Removing participants who indicated awareness of the S-C phase structure did not alter the results.

Discussion

Although cognitive control has traditionally been viewed in opposition to associative learning, recent studies have shown that people can learn to associate particular stimuli with specific attentional control states (e.g., high attentional focus). In Experiments 1 and 3, we examined whether these learned stimulus-control associations could transfer across pre-learned stimulus associations. We found evidence of a transfer effect, and in Experiment 2, showed that transfer of control state associations depended on the initial associations linking the stimuli pairs.

This transfer of learned control settings across associated stimuli represents a novel generalization mechanism in the domain of cognitive control. Our findings add to previous studies that have found evidence of context-specific control settings generalizing to new or unbiased stimuli (Crump & Milliken, 2009; Surrey et al., 2017; Weidler & Bugg, 2016; Wühr, Duthoo, & Notebaert, 2015) and members within particular social categories (Cañadas et al., 2013). We demonstrate that stimulus-specific control generalizes to arbitrarily associated stimuli: like reward (Wimmer & Shohamy, 2012), attentional control states can transfer across previously linked cues. Furthermore, this mechanism need not depend on an explicit understanding of the predictive relationships. In line with research on implicit context-control (Abrahamse, Braem, Notebaert, & Verguts, 2016; Crump, Vaquero, & Milliken, 2008), we show that people can learn to recruit high attentional focus in a flexible context-dependent manner, modulating their response to demands from closely linked contexts without explicit awareness of task structure.

An alternative explanation of our effects could be that participants modulated their response to Stroop stimuli through learned cue-response or cue-stimulus-response contingencies (Schmidt & Besner, 2008). However, we found no evidence of this learning in the Experiment 1 S-C phase data (see

Supplementary Material) or in the Experiment 3 replication. In the latter, we clearly demonstrated context-control learning in the absence of contingency learning, replicating previous studies (e.g., Crump & Milliken, 2009; Surrey et al., 2017; Weidler & Bugg, 2016; Wühr et al., 2015), as well as showing that only control-learning (but not stimulus-response contingency learning) transferred across linked stimuli.

This implicit stimulus-control binding is consistent with the concept of “event files,” episodic memory ensembles of stimulus and response features (Hommel, 2004). A recently proposed extension of this concept (Egner, 2014) suggests that stimulus and contextual features can be bound to internal representations of attentional states and cognitive strategies. Such “integrative” event files imply that closely linked cues can be included in an episodic file with their paired associates, subject to how stimuli are learned in the task structure.

The task structure may have influenced the stimulus-control learning that we observed. As in the stimulus-reward transfer study (Wimmer & Shohamy, 2012), we linked stimuli pairs, and stimuli and attentional control states, in a causal chain (S1 → S2 → control state) structure. This structure may have caused interference in the stimulus representations by disrupting predictive regularities, since S2 images act as goal-relevant outcomes in the S-S phase but predictive cues in the S-C phase. Alternatively, learned stimulus-control associations may also transfer across linked cues through knowledge of the covariational information provided by a common-cause (S2 → S1, S2 → control state) structure (Penn & Povinelli, 2007). Here, the predictive structure is more consistent, since the cues or “causes” are held constant, potentially allowing for better learning. If participants can learn stimulus-control associations through a common-cause structure, this indicates that transfer effects are observed only implicitly through associative learning, because associative and causal learning are treated independently. However, if participants cannot learn through a common-cause structure, this implies that strategy or reasoning adjustments are possible. Future studies should address this important question concerning the effect of structure on stimulus-control learning.

Future work should also examine whether people can transfer these control settings across different members of associated categories (cf. Cañadas et al., 2013), using trial-unique stimuli to control for priming effects and subsequently test for incidental encoding differences in the retrieval of control states. Control settings and retrieval may be further modulated on an individual level, prompting the need to pinpoint the optimal and boundary conditions for obtaining control-learning effects, which were highly variable in Experiment 1. Previous work has shown that individual differences in working memory capacity (Hutchison, 2011) and reward responsiveness (for review, see Braver, 2012) may modulate contextual control effects, while design considerations such as cue awareness (Farooqui & Manly, 2015) and timing

(e.g., Bugg & Smallwood, 2016) may also affect transfer learning. How exactly these various factors shape control-learning is yet to be fully understood.

In sum, we found that people can learn to associate specific stimuli with a state of high attentional selectivity and that these associations can transfer across paired-associates. This form of transfer learning has previously been demonstrated for simple actions (stimulus-response associations) and for reward associations, but to the best of our knowledge, the present study provides the first ever evidence for the associative transfer of stimulus-control associations across arbitrarily linked stimuli. This learning mechanism may form the basis of the human ability to generalize cognitive strategies over related contexts.

Author Note This research was supported in part by NIMH R01 MH 087610. All authors report no conflicts of interest.

References

- Abrahamse, E., Braem, S., Notebaert, W., & Verguts, T. (2016). Grounding cognitive control in associative learning. *Psychological Bulletin*, 142(7), 693–728.
- Appelbaum, L. G., Meyerhoff, K. L., & Woldorff, M. G. (2009). Priming and backward influences in the human brain: Processing interactions during the Stroop interference effect. *Cerebral Cortex*, 19(11), 2508–2521.
- Botvinick, M. M., Braver, T. S., Barch, D. M., Carter, C. S., & Cohen, J. D. (2001). Conflict monitoring and cognitive control. *Psychological Review*, 108(3), 624–652.
- Braver, T. S. (2012). The variable nature of cognitive control: A dual mechanisms framework. *Trends in Cognitive Sciences*, 16(2), 106–113.
- Bugg, J. M. (2017). Context, Conflict, and Control. In: T. Egner (Ed.), *The Wiley Handbook of Cognitive Control*, pp. 79–96. Oxford: Wiley-Blackwell.
- Bugg, J. M., & Crump, M. J. (2012). In support of a distinction between voluntary and stimulus-driven control: A review of the literature on proportion congruent effects. *Frontiers in Psychology*, 3, 367.
- Bugg, J. M., & Smallwood, A. (2016). The next trial will be conflicting! Effects of explicit congruency pre-cues on cognitive control. *Psychological Research*, 80(1), 16–33.
- Cañadas, E., Rodríguez-Bailón, R., Milliken, B., & Lupiáñez, J. (2013). Social categories as a context for the allocation of attentional control. *Journal of Experimental Psychology: General*, 142(3), 934–943.
- Crump, M. J., & Milliken, B. (2009). The flexibility of context-specific control: Evidence for context-driven generalization of item-specific control settings. *The Quarterly Journal of Experimental Psychology*, 62(8), 1523–1532.
- Crump, M. J., Vaquero, J. M., & Milliken, B. (2008). Context-specific learning and control: The roles of awareness, task relevance, and relative salience. *Consciousness and Cognition*, 17(1), 22–36.
- Egner, T. (2014). Creatures of habit (and control): A multi-level learning perspective on the modulation of congruency effects. *Frontiers in Psychology*, 5, 1247.
- Farooqui, A. A., & Manly, T. (2015). Anticipatory control through associative learning of subliminal relations: Invisible may be better than visible. *Psychological Science*, 26(3), 325–334.
- Hommel, B. (2004). Event files: Feature binding in and across perception and action. *Trends in Cognitive Sciences*, 8(11), 494–500.
- Hutchison, K. A. (2011). The interactive effects of listwide control, item-based control, and working memory capacity on Stroop

- performance. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 37(4), 851-860.
- Kanade, T., Cohn, J. F., & Tian, Y. (2000). Comprehensive database for facial expression analysis. In *Automatic Face and Gesture Recognition, 2000. Proceedings. Fourth IEEE International Conference on* (pp. 46-53). IEEE.
- King, J. A., Korb, F. M., & Egner, T. (2012). Priming of control: Implicit contextual cuing of top-down attentional set. *Journal of Neuroscience*, 32(24), 8192-8200.
- Le Pelley, M., & McLaren, I. (2003). Learned associability and associative change in human causal learning. *The Quarterly Journal of Experimental Psychology*, 56(1), 68-79.
- Norman DA, Shallice T (1986) Attention to action: Willed and automatic control of behavior. In: *Consciousness and Self-regulation: Advances in Research and Theory* (Davidson RJ, Schwartz GE, Shapiro D, eds), vol 4, pp 1–18. New York: Plenum Press.
- Penn, D. C., & Povinelli, D. J. (2007). Causal cognition in human and nonhuman animals: A comparative, critical review. *Annual Review of Psychology*, 58, 97-118.
- Schmidt, J. R., & Besner, D. (2008). The Stroop effect: Why proportion congruent has nothing to do with congruency and everything to do with contingency. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 34(3), 514-523.
- Stroop, J. R. (1935). Studies of interference in serial verbal reactions. *Journal of Experimental Psychology*, 18(6), 643-662.
- Surrey, C., Dreisbach, G., & Fischer, R. (2017). Context-specific adjustment of cognitive control: Transfer of adaptive control sets. *The Quarterly Journal of Experimental Psychology*, 70(11), 2386-2401.
- Weidler, B. J., & Bugg, J. M. (2016). Transfer of location-specific control to untrained locations. *The Quarterly Journal of Experimental Psychology*, 69(11), 2202-2217.
- Wimmer, G. E., & Shohamy, D. (2012). Preference by association: How memory mechanisms in the hippocampus bias decisions. *Science*, 338(6104), 270-273.
- Wühr, P., Duthoo, W., & Notebaert, W. (2015). Generalizing attentional control across dimensions and tasks: Evidence from transfer of proportion-congruent effects. *The Quarterly Journal of Experimental Psychology*, 68(4), 779-801.