Social Cognitive and Affective Neuroscience Advance Access published May 31, 2016

OXFORD

Social Cognitive and Affective Neuroscience, 2016, 1–9

doi: 10.1093/scan/nsw061
Original article

Does inhibitory control training transfer?: behavioral and neural effects on an untrained emotion regulation task

Kathryn G. Beauchamp, Lauren E. Kahn, and Elliot T. Berkman

Department of Psychology, University of Oregon, Eugene, OR, USA

Correspondence should be addressed to Kathryn G. Beauchamp, Department of Psychology, University of Oregon, 1227 University of Oregon, Eugene, OR 97403-1227, USA. E-mail: kgbeau@uoregon.edu

Abstract

Inhibitory control (IC) is a critical neurocognitive skill for successfully navigating challenges across domains. Several studies have attempted to use training to improve neurocognitive skills such as IC, but few have found that training generalizes to performance on non-trained tasks. We used functional magnetic resonance imaging (fMRI) to investigate the effect of IC training on a related but untrained emotion regulation (ER) task with the goal of clarifying how training alters brain function and why its effects typically do not transfer across tasks. We suggest hypotheses for training-related changes in activation relevant to transfer effects: the strength model and several plausible alternatives (shifting priorities, stimulus-response automaticity, scaffolding). Sixty participants completed three weeks of IC training and underwent fMRI scanning before and after. The training produced pre- to post-training changes in neural activation during the ER task in the absence of behavioral changes. Specifically, individuals in the training group demonstrated reduced activation during ER in the left inferior frontal gyrus and supramarginal gyrus, key regions in the IC neural network. This result is less consistent with the strength model and more consistent with a motivational account. Implications for future work aiming to further pinpoint mechanisms of training transfer are discussed.

Key words: inhibitory control; cognitive training; training transfer; functional neuroimaging; emotion regulation; strength (resource) model

Introduction

Inhibitory control (IC) is a critical neurocognitive skill for navigating cognitive, social and emotional challenges, and deficits in IC are a hallmark of some psychopathology (e.g. substance abuse; Perry and Carroll, 2008). IC is considered a basic element of the broader construct of self-control, which predicts positive outcomes such as academic achievement, relationship success (Tangney *et al.* , 2004; Duckworth, 2011), and substance use cessation (Berkman *et al.*, 2011; Mahmood *et al.*, 2012). As such, training IC is a promising intervention strategy. A critical open question in the emerging IC training literature is whether and how targeted IC training may transfer to conceptually related but untrained tasks.

Cognitive training studies to date have focused primarily on working memory (WM) and have shown mixed evidence for WM improvement through training (for review see Shipstead *et al.* , 2012). Though some of these studies have provided evidence of training transfer (Jaeggi *et al.*, 2011), many have not (for review, see Melby-Lervåg and Hulme, 2013). A recent study by Schweizer *et al.* (2013) showed some evidence of transfer from emotional WM training to both improved performance on and functional brain activation during an untrained ER task, suggesting a shared emotional component as a potential transfer mechanism. Even fewer studies have focused on IC training. Though IC performance appears to improve with training (Berkman *et al.*, 2014), evidence for training transfer has not yet

Received: 9 July 2015; **Revised:** 18 April 2016; **Accepted:** 25 April 2016

been found (Thorell *et al.*, 2009; Enge *et al.*, 2014). A recent meta-analysis of self-control training effects found them to be very small if they exist at all (Berkman, in press). Additionally, a well-powered recent study determined that playing video games, an activity that involves IC for certain types of games that were included in the analysis, did not improve performance on cognitive tasks (Unsworth *et al.*, 2015).

In light of these results, the field now needs a model of *how* IC training might transfer to untrained tasks. The overarching goal of the current paper is to provide a mechanistic framework that explains how IC training operates and, in so doing, accounts for the absence of transfer effects and suggests ways to improve them.

IC as a component of self-control

Inhibitory control (IC) is a key building block in the broader construct of self-control, or the ability to direct behavior away from short-term responses that conflict with long-term goals. This conceptualization of IC is consistent with the strength (or resource) model wherein self-control is a shared, limited resource that contributes to functioning across several response domains (Muraven and Baumeister, 2000). Though IC most closely maps on to the behavioral domain, it is nonetheless expected to be central to performance in other domains. For example, the abilities to inhibit motor and emotional responses are presumably related because they draw upon a shared overarching capacity for self-control. By the same logic, a training intervention that improves performance in one domain of self-control should improve performance in other domains.

The neural systems of IC

The neural systems activated during IC are fairly well characterized (Swick *et al.*, 2011; Wiecki & Frank, 2013). Successful IC typically engages the bilateral ventrolateral prefrontal cortex [VLPFC; primarily right inferior frontal gyrus (rIFG), but also the left], pre-supplementary motor area (preSMA), anterior cingulate cortex (ACC; Garavan *et al*, 2002) and the subthalamic nucleus (STN; Aron and Poldrack, 2006). Within this IC neural network, rIFG plays a critical role. Lesion work in both humans (Aron *et al.*, 2003) and non-human primates (Iversen and Mishkin, 1970) suggests that an intact rIFG is necessary for successful IC execution.

Neural overlap of IC and ER

Evidence from structural and functional neuroimaging supports the conceptualization of a unified self-control construct, with rIFG as a key component (Cohen and Lieberman, 2010). Individuals with rIFG lesions show impairments in both IC [i.e. as indexed by the Stop Signal Task (SST); Aron *et al.*, 2003] and ER abilities (i.e. as indexed by increased risk for affective psychopathology; Grafman *et al.*, 1986). Additionally, integrity of the right pars opercularis (a region within the rIFG) is associated with performance on both IC and ER tasks such that individuals with reduced gray matter intensity in this region perform worse on both tasks (Tabibnia *et al.*, 2011). Functional activation studies have consistently demonstrated rIFG activation during ER (Ochsner *et al.*, 2002; Wager *et al.* , 2008; Payer *et al.*, 2012), and those in which the same individuals complete separate ER and IC tasks show rIFG activation across both (Tabibnia *et al.*, 2011; Tabibnia *et al.*, 2014).

IC training

As a fundamental neurocognitive skill with a well-characterized neural network that is predictive of important outcomes and theoretically linked to other domains of self-control, IC is a highly promising target for intervention (Berkman *et al.*, 2012). Though IC appears to be trainable using an adaptive, standardized IC task (Berkman *et al.*, 2014; but see also Enge *et al.*, 2014), transfer effects of IC training have been elusive.

The few studies that have attempted to train IC have produced mixed results (Thorell *et al.*, 2009; Enge *et al.*, 2014; Berkman *et al.*, 2014; for review, Spierer *et al.*, 2013). A recent study using both Go/No-Go and Stop Signal Task (SST) training paradigms did not find behavioral evidence for training effects nor transfer to measures of fluid intelligence (Enge *et al.*, 2014). Notably, this study did not include any neural measures, precluding investigation of neural effects in the absence of behavioral differences. Another recent study from our lab demonstrated that a brief 3-week training on the SST does lead to improvement in IC that in turn related to differences in activation in the IC network (Berkman *et al.*, 2014). Specifically, the group that received the IC training showed an increase in activation in the rIFG from pre- to post-training during the cue phase of the task (i.e. before the use of IC), while the group that completed an active control task showed the opposite pattern. The fact that behavioral improvement on the SST was associated with activation changes in rIFG suggests that training with the SST successfully targeted the IC neural network.

Theoretical mechanisms of IC training transfer effects

Assuming that an IC training paradigm can successfully train IC, can this training transfer to an untrained ER task? Given the evidence that IC and ER share a common neural substrate alterable by IC training, IC training may transfer to ER by affecting this shared network.

The strength model makes distinct predictions about the behavioral and neural changes that might accompany training, thereby allowing for the identification of testable mechanisms by which transfer may or may not occur. The strength model proposes a strengthening (akin to the effect of strength training on muscles) of the common self-control resource through training. By practicing IC repeatedly, one is building the self-control 'muscle,' presumably making it larger and more robust. This strengthening is predicted to be task-general; any task requiring the self-control 'muscle' should benefit from its training.

The logic of most training studies follows the predictions of the strength model, either explicitly or implicitly: practice causes improvement by growing a shared resource. This popularity is likely attributable to the model's clear predictions regarding training effects and its strong foundation in the literature. However, there exist other plausible mechanisms for how training and transfer might work (Table 1). One such alternative is a motivational account of self-control, which has been described as a revision to the strength model. This motivational account, termed the 'shifting priorities' model, posits that a physiological resource is not necessary to explain the mechanisms of self-control. Instead, self-control failure is caused by a motivational/attentional switch from prioritizing 'have-to' tasks to prioritizing 'want-to' tasks (Inzlicht *et al.*, 2014). In this model, training would somehow increase motivation to work on an otherwise difficult, 'have-to' task, but it is unclear what the mechanism would be. Nonetheless, if motivation were increased—effectively transforming the 'have-to' task into a 'want-to' task—then the pattern of neural activity during the

training task might be characterized by task engagement (e.g. reduced default network activity; Anticevic *et al*., 2012).

Another plausible mechanism is that training might generate stimulus-response associations that, through automaticity, reduce the effortful cognitive burden of IC detection and implementation (Lenartowicz *et al*., 2011). In this case, we expect a pattern of improved behavioral performance in tandem with reduced activity in regions associated with effortful processing. Additionally, if training operates through stimulus-response pairing, then training effects are likely to be stimulus-specific; training-related gains in performance will not generalize to novel tasks that do not have the same stimuli and contingency structure as the training tasks.

A third alternative is that training causes a qualitative change in mental processing (i.e. use of different strategies). This prediction is related to the Scaffolding Theory of Aging and Cognition—revised, which posits that cognitive training in older adults may improve functioning by increasing neural activity in compensatory regions, rather than those underlying the trained cognitive skill per se (Reuter-Lorenz and Park, 2014). Though this theory was developed with older adult populations, the concept of scaffolding and its application to training is relevant across the lifespan. Specifically, scaffolding predicts that behavior change would be accompanied by neural activation in regions not otherwise recruited by the task.

The current study

The goal of this study is to investigate whether IC training transfers from a non-affective training task to an untrained affective task and to use neuroimaging to characterize the mechanism of transfer (or non-transfer). In a previous report (Berkman *et al*., 2014), we characterized the effect of a 3-week IC training program on functional neural activation *during IC*. However, as noted above, no study has yet demonstrated transfer of IC training effects to related yet untrained domains. Here, we report new analyses from the same dataset regarding the effects of the IC training on neural activation *during ER*.

Given previous mixed evidence of behavioral transfer effects, we do not have strong predictions about the improvement of behavioral ER capacity with IC training. However, we do predict changes in IC neural networks during the ER task as a result of IC training, specifically in the rIFG and possibly other regions that will provide evidence for or against the plausible mechanisms of training transfer described above.

We posit a set of competing hypotheses for patterns of brain activation with respect to models of training transfer outlined above. One possibility is that we will find evidence for an increase in activation in the IC neural network (specifically rIFG) consistent with strengthening of the self-control 'muscle' through training (i.e. increases in rIFG activation in training compared to sham). However, there are several alternative possibilities, whereby training will lead to a different pattern of neural results more consistent with one of the alternative models listed in Table 1.

Table 1. Plausible mechanisms of training transfer

	IC network activation	Other activation
Strength model	Increase	No change
Shifting priorities model	Decrease	Decrease in default mode network
Stimulus-response automaticity	Decrease	No change
Scaffolding	No change	Increase in compensatory regions

Materials and methods

Participants

Sixty participants (33 females, 27 males) aged 18–30 years (M = 21.63, s.d. = 2.99) were recruited through flyers posted around the University of Oregon (UO). The ethnic makeup of this group of participants was representative of the local community: 84% Caucasian, 4% Asian or Pacific Islander, 7% Hispanic and 5% other. Interested participants were screened by phone for eligibility (i.e. right-handedness, absence of neurological and mood disorders, absence of MRI contraindications). Eligible participants were scheduled for a baseline fMRI session at the UO's Lewis Center for Neuroimaging (LCNI). At the beginning of this session, all participants provided informed consent in accordance with a protocol approved by the UO institutional review board.

Procedure

Baseline, training and endpoint sessions occurred over the course of approximately 23 days for each participant. At baseline, participants underwent fMRI scanning at the LCNI during which they completed two runs of the SST (Verbruggen and Logan, 2008) followed by two runs of a cognitive reappraisal ER task (Gross, 1998). Participants also completed questionnaire measures following the scan that were unrelated to this study. At the end of the baseline session, participants were randomly assigned to either a training group or a sham-training group. Approximately 1–2 days (M = 1.58 days, s.d. = 0.72) following the baseline session, the training component began, consisting of 10 sessions of either the SST (training group) or a two-alternative forced-choice reaction time task (i.e. the SST without stop cues; sham group). Training sessions occurred approximately every other day for 3 weeks (M = 18.98 days, s.d. = 1.94) and were conducted in behavioral testing rooms in the Social and Affective Neuroscience (SAN) Laboratory in the Department of Psychology. All participants completed all 10 training sessions. For the endpoint session, participants returned to the LCNI approximately 1–2 days following their last training session for an fMRI scanning session identical to the baseline session except that the images used in the ER task were novel.

Tasks

Stop Signal. Each trial of the SST was comprised of a cue indicating the start of a trial (500 ms), then a go signal (1000 ms) consisting of an arrow pointing either left or right (with a ratio of 1:1), followed by an inter-trial interval of variable duration (M = 1400 ms; jittered following a gamma distribution). Participants were instructed to press the corresponding arrow key in response to the go signal. On a minority of trials (25%), an auditory stop signal played after the go signal at a variable latency called the stop-signal delay (SSD). On these trials, participants were instructed to inhibit their button press. The SSD was dynamically adjusted by 50 ms after each stop trial using a staircase function (i.e. increased following successful stops, decreased following failed stops). A rate of 50% response accuracy was achieved on stop trials through the alternating control of two independent

staircases over the SSD in blocks of eight trials. The difference between the speed of the stop process and the SSD is used to calculate the stop-signal response time (SSRT), the primary dependent measure of the SST. The integration method, which is less biased than the alternative mean method (Verbruggen *et al.*, 2013), was used to estimate the speed of the stop process. The SSRT was computed separately for each of the two runs of the SST and then averaged across the two runs for each time point (baseline, endpoint). Each run of the SST consisted of 128 trials (32 stop trials) and averaged 6:06 min in duration.

Emotion regulation. The cognitive reappraisal task used in this study was an adaptation of a task commonly used to assess ER (Gross, 1998). Each run of the task consisted of six blocks of five trials each. At the beginning of each block, participants were given one of three instructions (3000 ms): 'look and let yourself respond naturally' (Look), 'decrease emotion' (Reappraise) or 'scene description' (Label). The analyses reported here focus on the Look and Reappraise conditions because the neural regions supporting reappraisal are better understood (Ochsner *et al.*, 2012). The instruction screen was followed by a neutral or negative International Affective Picture System (IAPS) picture (5000 ms), a blank screen (500 ms) and a prompt for participants to rate their distress on a Likert scale of 0–9 (4500 ms). Each trial was separated by an inter-trial interval of variable duration (M ITI 1.85s). Each run of the ER task consisted of 30 trials and lasted 6:12 min.

Training sessions

During each training session, the training group completed one run of the SST that was modified as follows. First, the SSD from the previous session was used at the beginning of the subsequent training session to allow for continuous adaptation of difficulty level to maintain 50% accuracy across training sessions. Additionally, a strategy involving the cue at the start of each trial was employed to discourage participants from slowing their response to prepare for stop trials and instead encourage responding to go trials as quickly as possible. This strategy involved the start cue changing color (from white) based on the response time on the previous trial (to orange if previous response time >500 ms, to red if previous response time >750 ms). This procedure has been shown to decrease the bias and increase the efficiency of SSRT estimation (Verbruggen *et al.*, 2013). The sham group completed one run of the same task without sound.

Behavioral data analysis

Behavioral measures of interest included ER scores calculated from the ER task distress ratings (i.e. the average distress rating on Look trials minus that on Reappraise trials averaged across both runs; greater scores indicated more effective ER), and SSRTs estimated from the SST training sessions. We investigated the presence of behavioral training transfer from IC to ER scores by testing the group (sham *vs* training) × time (baseline *vs* endpoint) interaction. Within the training group, the slope of SSRTs across training sessions was calculated as a behavioral change measure to be incorporated into neuroimaging analyses. Outliers ($n = 2$) greater than 3s.d.'s above or below the mean were winsorized.

Imaging data acquisition

Neuroimaging data were acquired using a 3.0 T Siemens Allegra head-only scanner at the UO's LCNI. Data acquisition and preprocessing parameters were identical to those used by Berkman *et al.* (2014).

Statistical analyses were implemented in SPM8. For each subject, event-related condition effects were estimated according to the general linear model using a canonical hemodynamic response function, high-pass filtering (128 s) and a first-order autoregressive error structure. At the subject level, BOLD signal was modeled in a fixed effects analysis with regressors for negative and neutral Look trials, negative and neutral Reappraise trials, negative and neutral Label trials, and the instruction, cue and response periods totaling nine substantive regressors per run. Linear contrasts were created for each comparison of interest: Reappraise negative > Look negative at baseline vs. endpoint; Reappraise negative (> implicit baseline) at baseline *vs* endpoint; Reappraise negative (> implicit baseline) at baseline; and Reappraise negative (> implicit baseline) at endpoint. We included contrasts using the low-level baseline condition in order to reduce the number of factors present in our models and to minimize Type II error rates given that our analyses were not confirmatory in nature. These contrasts were then imported to group-level random effects analyses for inference to the population. Paired and independent samples *t*-tests were used to interrogate simple effects.

Neural IC training transfer effects to the ER task were measured in two ways: a group (sham *vs* training) × time (baseline *vs* endpoint) interaction and within training group analyses (correlated change from baseline to endpoint) incorporating behavioral change indices (i.e. IC training slope) as regressors. The logic of this latter analysis was to explore the extent to which training-related changes in neural activity during ER within the training group may relate to the quality of training. A combined voxel-height ($P = 0.005$) and cluster-extent ($k = 67$ for group × time interaction analyses; $k = 53$ for correlated change analyses) correction was applied for all analyses using Analysis of Functional Neuroimages AlphaSim software (Cox, 1996).

Results

Behavioral results: ER success

To verify that the ER task had the intended effects, paired samples t-tests were used to compare average distress ratings on Look trials to those on Reappraise trials. Participants had higher distress ratings on Look trials compared to Reappraise trials at both baseline, $t(59) = 14.82$, $P < 0.001$, and endpoint, $t(59) = 15.04$, $P < 0.001$, suggesting that participants reappraised successfully (Figure 1).

Behavioral results: training-induced change in ER performance

The extent to which training in IC transferred to ER performance was examined by testing the group × time interaction for ER scores from baseline to endpoint. The group × time interaction was not significant, indicating no difference in the ER score change from baseline to endpoint between the training and sham groups, $F(1,58) = 0.16$, *ns* (Figure 2). In other words, there was no behavioral training transfer from IC to ER.

Neuroimaging results: change in ER task-related neural activity as a function of training

Group × time interaction. To determine the extent to which neural activity during the ER task changed as a function of IC training, whole-brain analyses were conducted using the standard

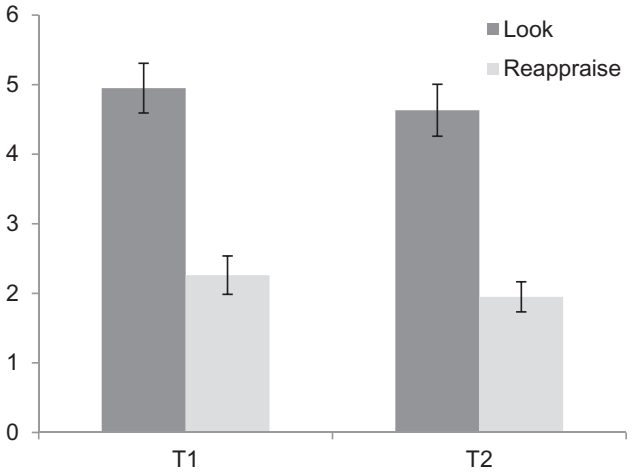

Fig. 1. Average distress ratings following Look and Reappraise conditions across time. Across groups, distress ratings for Reappraise trials were significantly less than those on Look trials at both baseline and endpoint, indicating successful ER.

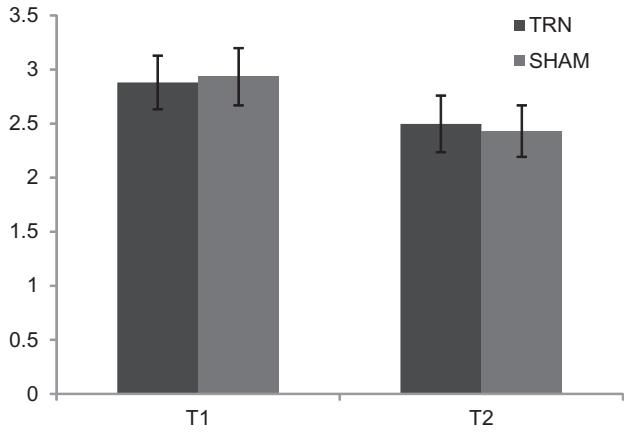

Fig. 2. Average ER score for training (TRN) and sham groups across time. The difference between baseline (T1) and endpoint (T2) ER scores was not significantly different between groups.

Reappraise > Look contrast. No regions demonstrated a significant group × time interaction for the Reappraise > Look contrast. When the group × time interaction was investigated during the Reappraise period (>implicit baseline), four significant clusters emerged (Table 2), including one in the left IFG (−45, 21, 18; Figure 3), a part of the IC network, and one in the supramarginal gyrus (36, −60, 39), in which the sham group showed a greater increase in these regions from baseline to endpoint than did the training group. Parameter estimates were extracted to interrogate the group × time interaction further Figure 4. Simple effects tests revealed that in three of the four regions, the interaction was driven by both a significant increase in activation over time in the sham group [left occipital, $t(29) = -2.768$, $P = 0.01$; supramarginal gyrus, $t(29) = -2.89$, $P = 0.007$; right occipital, $t(29) = -2.67$, $P = 0.012$] and also a significant decrease in activation over time in the training group [left occipital, $t(29) = 3.187$, $P = 0.003$; supramarginal gyrus, $t(29) = 2.74$, $P = 0.011$; right occipital, $t(29) = 3.27$, $P = 0.003$] (Figure 4). In the left IFG, activation significantly decreased from baseline to endpoint in the training group, $t(29) = 4.85$, $P < 0.001$, but not in the sham group, $t(29) = -1.29$, *ns*. Simple effects tests also revealed significantly higher activation at baseline in the training group compared to the sham group in the supramarginal gyrus, $t(58) = -2.475$, $P = 0.016$, and significantly lower activation in the training group at endpoint compared to the sham group in all clusters [left occipital, $t(58) = 2.81$, $P = 0.026$; left IFG, $t(58) = 3.24$, $P = 0.002$; supramarginal gyrus, $t(58) = 3.24$, $P = 0.002$; right occipital, $t(58) = 3.03$, $P = 0.004$].

Neuroimaging results: activations correlated with IC improvement in training group across time

We used a 'correlated change' model to estimate the degree of association between improvements in IC performance over time (training slope) and changes in activation during ER (>implicit baseline) from baseline to endpoint. Positive correlated change reflected regions where changes in activation during ER from pre- to post-training were linearly related to changes in behavioral improvement in IC, independent of their pre- or post-training level. This analysis revealed a significant association between change in activation from baseline to endpoint in the medial prefrontal cortex (mPFC, a component of the default mode network) and training slope such that a more negative training slope (i.e. better IC performance over time) was associated with decreased activation in this area across time (Figure 5; middle). Simple effects demonstrated that greater deactivation from baseline to endpoint in this region was associated with a more negative training slope (Figure 5, right). Coordinates of all clusters significantly correlated with training slope are presented in Table 3.

To further clarify the correlated change result in the mPFC, we examined the relationship of activation in mPFC with activation in the previously identified left IFG cluster. These parameter estimates were significantly correlated such that decreases in mPFC activation over time were associated with decreases in left IFG activation in the training group [$r(28) = 0.479$, $P < 0.01$]. We discuss possible interpretations of this relationship in light of the outlined plausible mechanisms of training transfer below.

Discussion

This study aimed to examine whether IC training leads to changes in an untrained ER task. Transfer of behavioral improvement to untrained tasks has been elusive in cognitive training studies. A more refined understanding of how training alters specific neurocognitive systems may explain the lack of transfer across presumably related domains. Here, we investigated the effects of a training paradigm known to engage the IC neural network on functional activation during an ER task. Our analyses replicated the previously observed lack of behavioral transfer and provided some potentially diagnostic insights based on patterns of change in brain activation. Specifically, our results indicate that IC training may lead to increased task engagement and decreased default mode network activity (e.g. mPFC) during ER, which is most consistent with a motivational account of self-control among the possibilities we considered.

The training and sham groups demonstrated comparable performance on the ER task both at baseline and endpoint, indicating no behavioral effect of IC training on ER performance. This result is inconsistent with the strength model, which predicts that self-control training in any domain should lead to domain-general improvements. Berkman *et al.* (2014) found evidence that IC training effects in the task deployed here were critically tied to the cue used in training, consistent with

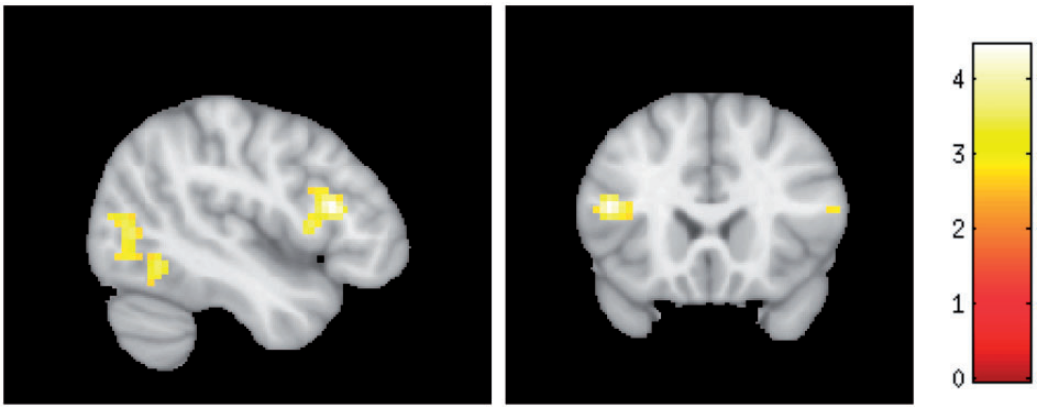

Fig. 3. Group × time interaction in the left IFG. The left IFG showed greater activation during Reappraise trials (>implicit baseline) from baseline to endpoint in the sham group compared to the training group. Corrected using AlphaSim; voxelwise threshold of $P = 0.005$, $k > 67$.

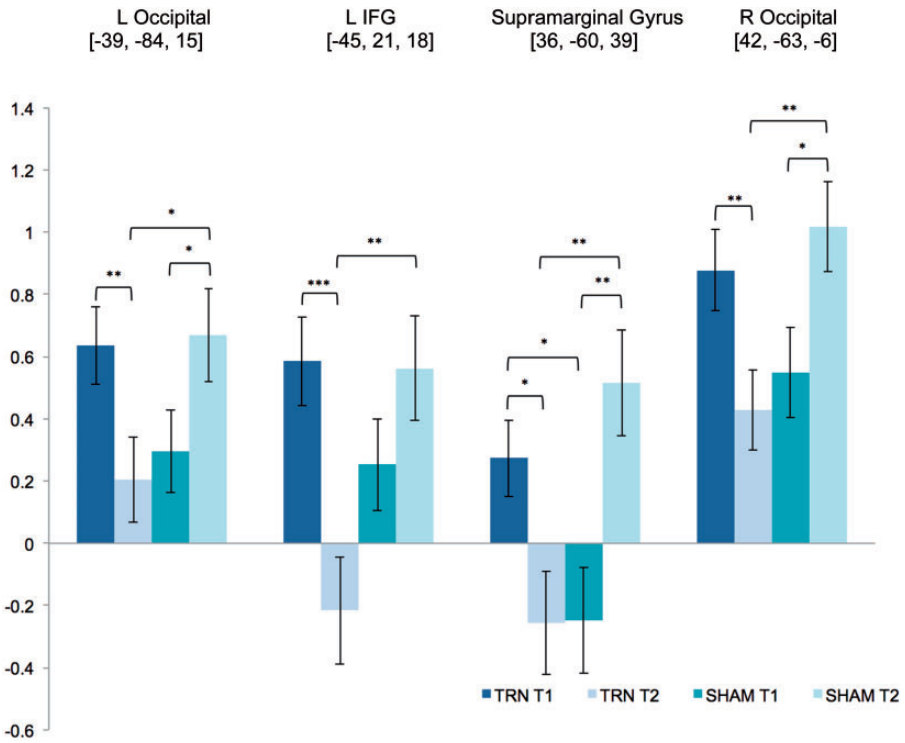

Fig. 4. Parameter estimates for significant clusters in the group × time interaction for the contrast Reappraise > implicit baseline; *P < 0.05, **P < 0.01, ***P < 0.001.

Table 2. Reappraise > implicit baseline; group–time interaction

	Anatomical region	x	y	z	k	t	z
T2 > T1 for Training > Sham	–	–	–	–	–	–	–
T2 > T1 for Sham > Training	L occipital	−39	−84	15	105	3.49	3.31
	lIFG	−45	21	18	236	4.44	4.10
	Supramarginal gyrus	36	−60	39	155	4.21	3.92
	R occipital	42	−63	−6	75	4.06	3.79

Corrected using AlphaSim; voxelwise threshold of $P = 0.005$, $k > 67$.

implicit learning models that suggest performance improvements become specific to sets of cues that are paired with training trials (Lenartowicz *et al.*, 2011). In this case, the fact that our participants saw different cues in trained and untrained tasks could account for a lack of behavioral transfer effects.

Despite the lack of behavioral effects, brain activation differences during the ER task did emerge as a result of IC training. A significant decrease in activation in the training group (*vs* sham group) was found in several regions including the left IFG and supramarginal gyrus, both regions associated with training on

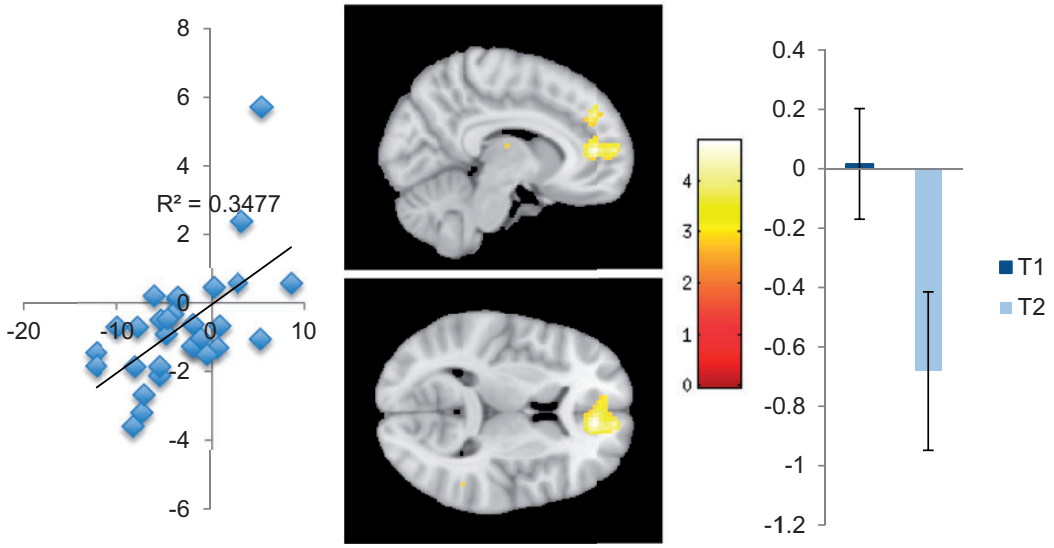

Fig. 5. Correlation with IC improvement in the mPFC in the training group across time (i.e. activation in Reappraise > implicit baseline contrast from T1 to T2 positively correlated with training slope; brain images, middle). Activation in the mPFC during Reappraise trials at endpoint (T2) compared to baseline (T1) was positively correlated with training slope [scatter plot, left; decreased mPFC activation associated with more negative training slope (better IC)]. Parameter estimates from the simple effects at each timepoint (right) demonstrated an increase in deactivation in the mPFC from baseline to endpoint during Reappraise trials. Corrected using AlphaSim; voxelwise threshold of $P = 0.005$, $k > 53$.

Table 3. Reappraise > implicit baseline; correlations with IC improvement (i.e. slope of SSRTs across training sessions) in training group across time

	Anatomical region	x	y	z	k	t	z
Pos. correlated with IC improvement	cingulate	−12	0	42	57	3.72	3.33
	SMA	−21	−30	57	84	3.87	3.44
	dmPFC	15	45	36	67	4.09	3.59
	mPFC	9	45	9	228	4.77	4.05
Neg. correlated with IC improvement	–	–	–	–	–	–	–

Corrected using AlphaSim; voxelwise threshold of $P = 0.005$, $k > 53$.

the SST in previous work (Berkman *et al.*, 2014). Interestingly, the training group demonstrated a pattern of reduced activation in these regions during ER from baseline to endpoint, the opposite pattern of activation change observed by Berkman *et al.* (2014) in the training group during the cue phase of the SST (i.e. before IC). It is possible that this pattern of reduced activation represents an increase in the strength of stimulus-response associations related to regulating emotion such that those participants who trained on IC achieved similar behavioral ER to those who did not due to a reduction in the cognitive burden associated with ER through automaticity. An increase in neurocognitive efficiency could also explain these results; however, this explanation is speculative as additional data (i.e. information about connectivity between local and global neural networks) are needed to empirically determine the extent to which neural efficiency may differ between training and sham groups (Poldrack, 2014).

Analyses within the training group suggest that IC training might transfer through a motivational mechanism. The mPFC, a component of the default mode network, demonstrated marked decreases in activation during ER that were associated with increases in IC performance across training. This increase in deactivation in mPFC is typically associated with an increase in task engagement (Gusnard *et al.*, 2001; Raichle *et al.*, 2001). Thus, IC training may have affected neural activation during ER through decreases in default mode network activity that enabled increased engagement in the self-control task at hand. Intriguingly, parameter estimates of this mPFC activation were highly correlated with those of activation in the left IFG obtained from the group × time interaction model, suggesting that reductions in default mode network activity might contribute to more efficient processing in the IC network. This pattern of brain activation could be interpreted as evidence for both a motivational account as well as scaffolding in which activity in one network may account for changes in another network. This interpretation highlights the important point that the proposed mechanisms are not necessarily in opposition or mutually exclusive.

A key limitation of the current study is that the design of the ER task was not optimized to investigate other potential mechanisms of training transfer beyond the strength model. Based on the findings of Berkman *et al.* (2014), an ER task sharing a cue with the trained IC task would be most likely to demonstrate training transfer effects through a reactive to proactive shift in control (Braver, 2012). Future studies can test for training transfer effects by incorporating shared cues across trained and non-trained tasks and explicitly designing non-trained tasks to allow for interrogation of the cue phase of the trial (Denny *et al.*, 2014). Additionally, the use of a motivationally salient cue (e.g. a desired consumable) may increase the effectiveness of the IC training through the proposed motivational mechanism.

The absence of behavioral differences in the untrained task following training in the presence of brain activation differences is notable. It is possible that a larger dose (e.g. more frequent or longer sessions) would have facilitated behavioral training transfer. Additionally, a larger overall sample size would increase detection rates. However, a more refined approach for future neuroimaging studies would be to increase the sampling precision (e.g. by selecting individuals with IC deficits). Moderating factors, including IC ability at baseline, may account for the dissociation between behavioral and neural training transfer effects. Neural activation may also be a precursor to behavioral change. Future studies can incorporate these perspectives to inform theories of training and transfer effects.

Another limitation lies in the significant difference in the supramarginal gyrus activations during reappraisal at baseline between the training and sham groups. Though this pre-existing activation difference makes interpretation of training transfer effects more challenging, the use of random assignment to condition, lack of behavioral differences at baseline and analyses that incorporate change over time serve to strengthen the current results.

In sum, IC training may generalize to an untrained ER task initially at the level of the brain and perhaps eventually in behavioral performance. Our results provide some evidence for increased automaticity in processing in the left IFG, a key component of the IC network, during ER following training, a pattern of results inconsistent with the strength model that proposes an increase in a common self-control 'muscle' with training. Overall, our results support the existence of neural training transfer effects related to increases in automaticity in key IC neural network regions as well as decreases in default mode network activity associated with increased task engagement, pinpointing likely mechanisms of training transfer for future investigations.

Funding

This work was supported by the National Institutes of Health (AG048840, CA175241, DA035763 to E.T.B.).

References

Anticevic, A., Cole, M.W., Murray, J.D., Corlett, P.R., Wang, X.J., Krystal, J.H. (2012). The role of default network deactivation in cognition and disease. *Trends in Cognitive Sciences*, **16**(12), 584–92. http://doi.org/10.1016/j.tics.2012.10.008.

Aron, A.R., Fletcher, P.C., Bullmore, T., Sahakian, B.J., Robbins, T.W. (2003). Stop-signal inhibition disrupted by damage to right inferior frontal gyrus in humans. *Nature Neuroscience*, **6**(2), 115–6.

Aron, A.R., Poldrack, R.A. (2006). Cortical and subcortical contributions to stop signal response inhibition: role of the subthalamic nucleus. *The Journal of Neuroscience*, **26**(9), 2424–33.

Berkman, E.T., Falk, E.B., Lieberman, M.D. (2011). In the trenches of real-world self control neural correlates of breaking the link between craving and smoking. *Psychological Science*, **22**(4), 498–506.

Berkman, E.T., Graham, A.M., Fisher, P.A. (2012). Training self-control: a domain general translational neuroscience approach. *Child Development Perspectives*, **6**(4), 374–84.

Berkman, E.T., Kahn, L.E., Merchant, J.S. (2014). Training-induced changes in inhibitory control network activity. *The Journal of Neuroscience*, **34**(1), 149–57.

Berkman, E.T. (in press). Self-regulation training. In: Vohs, K.D., Baumeister, R.F., editors. *Handbook of Self-Regulation: Research, Theory and Applications*. New York: Guildford.

Braver, T.S. (2012). The variable nature of cognitive control: a dual mechanisms framework. *Trends in Cognitive Sciences*, **16**(2), 106–13.

Cohen, J.R., Lieberman, M.D. (2010). The common neural basis of exerting self-control in multiple domains. In: Trope, Y., Hassin, R., Ochsner, K.N., editors. *Self Control in Society, Mind, and Brain*. UK: Oxford University Press, 141–62.

Cox, R.W. (1996). AFNI: software for analysis and visualization of functional magnetic resonance neuroimages. *Computers and Biomedical Research*, **29**(3), 162–73.

Denny, B.T., Ochsner, K.N., Weber, J., Wager, T.D. (2014). Anticipatory brain activity predicts the success or failure of subsequent emotion regulation. *Social Cognitive and Affective Neuroscience*, **9**(4), 403–11.

Duckworth, A.L. (2011). The significance of self-control. *Proceedings of the National Academy of Sciences of the United States of America*, **108**(7), 2639–40.

Enge, S., Behnke, A., Fleischhauer, M., Küttler, L., Kliegel, M., Strobel, A. (2014). No evidence for true training and transfer effects after inhibitory control training in young healthy adults. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, **40**(4), 987–1001.

Garavan, H., Ross, T.J., Murphy, K., Roche, R.A.P., Stein, E.A. (2002). Dissociable executive functions in the dynamic control of behavior: inhibition, error detection, and correction. *NeuroImage*, **17**(4), 1820–9.

Grafman, J., Vance, S.C., Weingartner, H., Salazar, A.M., Amin, D. (1986). The effects of lateralized frontal lesions on mood regulation. *Brain*, **109**(6), 1127–48.

Gross, J.J. (1998). Antecedent-and response-focused emotion regulation: divergent consequences for experience, expression, and physiology. *Journal of Personality and Social Psychology*, **74**(1), 224.

Gusnard, D.A., Akbudak, E., Shulman, G.L., Raichle, M.E. (2001). Medial prefrontal cortex and self-referential mental activity: relation to a default mode of brain function. *Proceedings of the National Academy of Sciences of the United States of America*, **98**(7), 4259–64.

Inzlicht, M., Schmeichel, B.J., Macrae, C.N. (2014). Why self-control seems (but may not be) limited. *Trends in Cognitive Sciences*, **18**(3), 127–33.

Iversen, S.D., Mishkin, M. (1970). Perseverative interference in monkeys following selective lesions of the inferior prefrontal convexity. *Experimental Brain Research*, **11**(4), 376–86.

Jaeggi, S.M., Buschkuehl, M., Jonides, J., Shah, P. (2011). Short-and long-term benefits of cognitive training. *Proceedings of the National Academy of Sciences of the United States of America*, **108**(25), 10081–6.

Lenartowicz, A., Verbruggen, F., Logan, G.D., Poldrack, R.A. (2011). Inhibition-related activation in the right inferior frontal gyrus in the absence of inhibitory cues. *Journal of Cognitive Neuroscience*, **23**(11), 3388–99.

Mahmood, O.M., Goldenberg, D., Thayer, R., Migliorini, R., Simmons, A.N., Tapert, S.F. (2012). Adolescents' fMRI activation to a response inhibition task predicts future substance use. *Addictive Behaviors*, **38**(1), 1435–41

Melby-Lervåg, M., Hulme, C. (2013). Is working memory training effective? A meta-analytic review. *Developmental Psychology*, **49**(2), 270–91.

Muraven, M., Baumeister, R.F. (2000). Self-regulation and depletion of limited resources: [Database] does self-control resemble a muscle? *Psychological Bulletin*, **126**(2), 247–59.

Ochsner, K.N., Bunge, S.A., Gross, J.J., Gabrieli, J.D. (2002). Rethinking feelings: an fMRI study of the cognitive regulation of emotion. *Journal of Cognitive Neuroscience*, **14**(8), 1215–29.

Ochsner, K.N., Silvers, J.A., Buhle, J.T. (2012). Functional imaging studies of emotion regulation: a synthetic review and evolving model of the cognitive control of emotion. *Annals of the New York Academy of Sciences*, **1251**(1), E1–24.

Payer, D.E., Baicy, K., Lieberman, M.D., London, E.D. (2012). Overlapping neural substrates between intentional and incidental down-regulation of negative emotions. *Emotion*, **12**(2), 229–35.

Perry, J.L., Carroll, M.E. (2008). The role of impulsive behavior in drug abuse. *Psychopharmacology*, **200**(1), 1–26.

Poldrack, R.A. (2014). Is "efficiency" a useful concept in cognitive neuroscience? *Developmental Cognitive Neuroscience*, **11**, 12–7.

Raichle, M.E., MacLeod, A.M., Snyder, A.Z., Powers, W.J., Gusnard, D.A., Shulman, G. L., (2001). A default mode of brain function. *Proceedings of the National Academy of Sciences of the United States of America*, **98**(2), 676–82.

Reuter-Lorenz, P.A., Park, D.C. (2014). How does it STAC up? Revisiting the scaffolding theory of aging and cognition. *Neuropsychology Review*, **24**(3), 355–70.

Schweizer, S., Grahn, J., Hampshire, A., Mobbs, D., Dalgleish, T. (2013). Training the emotional brain: improving affective control through emotional working memory training. *The Journal of Neuroscience*, **33**(12), 5301–11.

Shipstead, Z., Redick, T.S., Engle, R.W. (2012). Is working memory training effective? *Psychological Bulletin*, **138**(4), 628–54.

Spierer, L., Chavan, C.F., Manuel, A.L. (2013). Training-induced behavioral and brain plasticity in inhibitory control. *Frontiers in Human Neuroscience* **7**, 427.

Swick, D., Ashley, V., Turken, U. (2011). Are the neural correlates of stopping and not going identical? Quantitative meta-analysis of two response inhibition tasks. *NeuroImage*, **56**(3), 1655–65.

Tabibnia, G., Monterosso, J.R., Baicy, K., *et al.* (2011). Different forms of self-control share a neurocognitive substrate. *The Journal of Neuroscience*, **31**(13), 4805–10.

Tabibnia, G., Creswell, J.D., Kraynak, T.E., Westbrook, C., Julson, E., Tindle, H.A. (2014). Common prefrontal regions activate during self-control of craving, emotion, and motor impulses in smokers. *Clinical Psychological Science*, **5**, e1000529.

Tangney, J.P., Baumeister, R.F., Boone, A.L. (2004). High self-control predicts good adjustment, less pathology, better grades, and interpersonal success. *Journal of Personality*, **72**(2), 271–324.

Thorell, L.B., Lindqvist, S., Bergman Nutley, S., Bohlin, G., Klingberg, T. (2009). Training and transfer effects of executive functions in preschool children. *Developmental Science*, **12**(1), 106–13.

Unsworth, N., Redick, T.S., McMillan, B.D., Hambrick, D.Z., Kane, M.J., Engle, R.W. (2015). Is playing video games related to cognitive abilities? *Psychological Science*, **26**(6), 759–74.

Verbruggen, F., Chambers, C.D., Logan, G.D. (2013). Fictitious inhibitory differences how skewness and slowing distort the estimation of stopping latencies. *Psychological Science*, **24**(3), 352–62.

Verbruggen, F., Logan, G.D. (2008). Response inhibition in the stop-signal paradigm. *Trends in Cognitive Sciences*, **12**(11), 418–24.

Wager, T.D., Davidson, M.L., Hughes, B.L., Lindquist, M.A., Ochsner, K.N. (2008). Prefrontal-subcortical pathways mediating successful emotion regulation. *Neuron*, **59**(6), 1037–50.

Wiecki, T.V., Frank, M.J. (2013). A computational model of inhibitory control in frontal cortex and basal ganglia. *Psychological Review*, **120**(2), 329.