NeuroImage 151 (2017) 92–104

Contents lists available at ScienceDirect

NeuroImage

journal homepage: www.elsevier.com/locate/neuroimage

Neural mechanisms of cue-approach training

Akram Bakkour[a,b,1], Jarrod A. Lewis-Peacock[a,b,c], Russell A. Poldrack[a,b,c,2], Tom Schonberg[a,d,*]

[a] Imaging Research Center, The University of Texas at Austin, 100 E 24th St, Stop R9975, Austin, TX 78712, USA
[b] Department of Neuroscience, The University of Texas at Austin, 100 E 24th St, Stop C7000, Austin, TX 78712, USA
[c] Department of Psychology, The University of Texas at Austin, 108 E Dean Keeton, Stop A8000, Austin, TX 78712, USA
[d] Department of Neurobiology, Faculty of Life Sciences and Sagol School of Neuroscience, Tel Aviv University, P.O. Box 39040, Tel Aviv 6997801, Israel

ARTICLE INFO

Keywords:
Cue-approach training
Behavioral change
Value-based decision making

ABSTRACT

Biasing choices may prove a useful way to implement behavior change. Previous work has shown that a simple training task (the cue-approach task), which does not rely on external reinforcement, can robustly influence choice behavior by biasing choice toward items that were targeted during training. In the current study, we replicate previous behavioral findings and explore the neural mechanisms underlying the shift in preferences following cue-approach training. Given recent successes in the development and application of machine learning techniques to task-based fMRI data, which have advanced understanding of the neural substrates of cognition, we sought to leverage the power of these techniques to better understand neural changes during cue-approach training that subsequently led to a shift in choice behavior. Contrary to our expectations, we found that machine learning techniques applied to fMRI data during non-reinforced training were unsuccessful in elucidating the neural mechanism underlying the behavioral effect. However, univariate analyses during training revealed that the relationship between BOLD and choices for Go items increases as training progresses compared to choices of NoGo items primarily in lateral prefrontal cortical areas. This new imaging finding suggests that preferences are shifted via differential engagement of task control networks that interact with value networks during cue-approach training.

1. Introduction

In order to eliminate unhealthy behaviors, one must find ways to enhance healthy choices. Changing preferences is an important strategy in addressing public health concerns, such as the obesity epidemic. To achieve lasting behavioral change to improve health, one must overcome the automaticity and strength of first-learned habits. First-learned behaviors are the rule that must be broken by subsequent learning in order for new habits to replace older ones over the long term (Bouton, 2004). Initial positive change in behavior may be achieved through intervention based on willful effort (Schonberg et al., 2014b; Tricomi et al., 2009), but the long term prospects for such improvement are uncertain (Bjork, 2001; Bouton, 1993; Cahill and Perera, 2011; Higgins et al., 1995; Wood and Neal, 2007). Focus has turned to targeting automatic processes to change human behavior with the goal of preventing disease (Marteau et al., 2012).

Previous research on value-based decision making has focused mostly on external reinforcement (O'Doherty et al., 2004; Thorndike, 1911) or the description of the decision problem (De Martino et al., 2006; Slovic, 1995; Tversky and Kahneman, 1986), but few have attempted to directly influence the underlying subjective values of individual options. In previous work by our group, we showed that choices can be biased toward targeted food items and the subjective value placed on these items can be differentially modulated by simply associating particular food items with an auditory cue to perform a motor response, without relying on external reinforcement or reframing the decision problem (Schonberg et al., 2014a). The previously described cue-approach task (CAT) is similar to the cued inhibition version of the stop-signal task (Lenartowicz et al., 2011; Verbruggen and Logan, 2008), with a crucial difference. In a typical stop-signal task, participants press a button on the keyboard every time a stimulus appears on the screen, except when a tone sounds they must try to inhibit a prepotent motor response. In CAT however, participants passively view stimuli on the screen, except when a tone sounds, they must press a button on the keyboard as quickly as possible. Training inhibition has been demonstrated to influence choice behavior for appetitive stimuli (Houben et al., 2012; Lawrence et al., 2015; Veling et al., 2013) and value for neutral stimuli (Wessel et al., 2014).

* Correspondence to: Department of Neurobiology, Faculty of Life Sciences and Sagol School of Neuroscience, Tel Aviv University, P.O. Box 39040, Tel Aviv 6997801, Israel.
E-mail address: schonberg@post.tau.ac.il (T. Schonberg).
[1] Present address: Department of Psychology, Columbia University, 1190 Amsterdam Ave, MC5501, New York, NY, 10027, USA.
[2] Present address: Department of Psychology, Stanford University, Bldg. 420, Jordan Hall, Stanford, CA, 94305, USA.

http://dx.doi.org/10.1016/j.neuroimage.2016.09.059
Accepted 23 September 2016
Available online 24 September 2016
1053-8119/

Following stop-signal or go/no-go inhibition, participants tended to avoid or devalue stimuli that were associated with inhibition of action. However, rather than aiming to decrease choices, we developed CAT seeking to enhance choices for certain stimuli. In the original version of CAT, participants were asked to fast for four hours prior to arriving for the experiment. After providing informed consent, they were endowed with $3 to take part in an auction to obtain their pre-experimental preferences for 60 food items (Becker et al., 1964; Plassmann et al., 2007). Items were then rank ordered based on preference and median split into high and low value items. High and low value items were then placed into one of two experimental conditions: Go or NoGo. During training, participants passively viewed pictures of food items and pressed a button when they heard an infrequent tone. In a subsequent probe phase, participants chose one item from a pair of equally preferred items, one associated with a tone during training (Go) and the other not associated with a tone (NoGo). Cue-approach training has proven to directly influence preference for single items through choice behavior following training. Approached (Go) items were chosen more often than initially equally preferred, non-approached (NoGo) items (Bakkour et al., 2016; Schonberg et al., 2014a). This procedure successfully changed choice behavior and the effect was maintained over six to eight weeks for participants who underwent the longest training (Schonberg et al., 2014a). Such a shift in choice behavior is thought to be mediated by an increase in gain in the coding of value for Go items in the ventromedial prefrontal cortex (vmPFC, Schonberg et al., 2014a), a brain region that has previously been heavily implicated in coding for value (Bartra et al., 2013; Padoa-Schioppa and Assad, 2006). This work has established cue-approach training as a model for non-reinforced preference change via modulation of subjective value for individual items. The question remains; how are values of Go items being modulated during CAT training?

Development of CAT was influenced by work on the attentional boost effect (Lin et al., 2010; Swallow and Jiang, 2010). In a typical attentional boost task, participants have better subsequent memory for incidental stimuli that were presented along with targets than those that were presented along with non-targets. The attentional boost effect established the importance of behavioral relevance in improving memory for incidental information. The cue-approach effect similarly established the importance of behavioral relevance for shifting preferences. Follow-up behavioral studies (Bakkour et al., 2016), using variations on the basic cue-approach training task have singled out memory retrieval and sustained top-down attention mechanisms to be at play during cue-approach training, leading to a shift in preferences at a later choice phase. However, standard univariate analyses of training-phase fMRI data in the previous imaging study of CAT were inconclusive and did not provide any insight into the neural mechanism responsible for modulating values of individual items during CAT training (Schonberg et al., 2014a). In the current study, we set out to characterize changes in neural activity during the cue-approach training phase using both univariate and multivariate analysis techniques.

Machine learning and pattern recognition algorithms have recently been adapted and developed to decode and characterize cognitive task-relevant neural activity using fMRI data (see Lemm et al., 2011; Mahmoudi et al., 2012, for review). One of the most popular of these machine-learning techniques is linear classification. This is a technique for decoding information about task variables from patterns of activity across an array of voxels. One of the common linear classification algorithms is the linear support vector machine (SVM). In this study, we sought to train a linear SVM classifier to identify whole-brain fMRI patterns elicited by cognitive processes thought to underlie shifts in choice preference during cue-approach training. Our hypothesis was that changes in classifier identification of the level of engagement of these cognitive processes of interest during training would predict later choices, reflecting a shift in preferences.

In order to test our hypothesis, we developed a cognitive localizer task that engages three distinct cognitive processes implicated in value change during the cue-approach training task: perception, memory retrieval, and valuation. We used multivariate pattern analysis techniques on fMRI data acquired during this novel task to predict the level of engagement of these cognitive processes during cue-approach training. We investigated how changes in these processes (as measured by classifier predictions) contributed to a shift in preferences at a later choice phase. This analysis allows us to directly test our hypothesis that changes in the level of engagement of these particular cognitive processes during training predicts a shift in choice behavior. We expect that increases in the engagement of valuation and memory retrieval processes over the course of cue-approach training will be related to later choices. Furthermore, we were able to test whether process engagement progressed differentially for Go and NoGo trials as training proceeded. We predicted that the differential change in engagement of valuation and memory retrieval processes from beginning to end of the training phase, rather than the difference in overall engagement of these processes, would be predictive of later choices as participants learn to associate the food item with the tone cue as the training phase progresses. This allowed us the potential to better understand the neural mechanisms underlying non-reinforced training that leads to a shift in preferences. Finally, we also used standard univariate fMRI analysis techniques on probe phase data to replicate previous findings, and on training phase data to identify changes in whole-brain activation throughout training.

The design of the current study was optimized for application of MVPA techniques to identify underlying neurocognitive mechanisms for the CAT effect. Previous studies have demonstrated the power of these techniques not only to classify distributed patterns of fMRI activity elicited by different categories of images while the participant was viewing them (Cox and Savoy, 2003; Haxby et al., 2001), but also to classify intentions (Haynes et al., 2007; Soon et al., 2008), attentional states (Rosenberg et al., 2015) and the contents of memory recall (Polyn et al., 2005) using classifiers trained on different sets of stimuli from those being classified. Furthermore, and most germane to our main question of interest in the current study, Gross et al. (2014) trained an SVM classifier to discriminate levels of subjective value of foods and predicted the subjective value of engaging activities and vice versa. This supports the idea of common representation of value and the valuation process across domains. This finding also suggests that classifying the valuation process in one task can be used to decode value from a different task as planned in the current study. Other studies demonstrated robust cross-modal or cross-task classification (Lewis-Peacock et al., 2012, 2015). Polyn et al. (2005) trained classifiers on fMRI data from a localizer task requiring the perception and evaluation of familiar pictures, and then used these classifiers to decode the category of stimuli being retrieved from long-term memory during free recall. Lewis-Peacock and Postle (2008) used the same localizer task and analysis approach to decode the contents of working memory during cued recall. Esterman et al. (2009) used fMRI pattern classifiers to decode which domain of cognitive control (e.g., shifting visuospatial attention, switching task rules, shifting attention in working memory) was engaged at any given moment. Together, these findings suggest that fMRI classifiers trained on long-term memory retrieval might be able to identify the engagement of this process during CAT training.

2. Materials and methods

2.1. Participants

Thirty-two healthy right-handed participants (17 female, mean age=21.8 ± 3.1, age range: 18–29, mean body mass index (BMI) =22.3 ± 3.8) completed the standard CAT while in a magnetic resonance imaging (MRI) scanner.

All participants had normal or corrected-to-normal vision, no history of psychiatric, neurologic or metabolic illnesses, no history of