NeuroImage 115 (2015) 214–223

Contents lists available at ScienceDirect

NeuroImage

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

Expectation modulates neural representations of valence throughout the human brain

Ashwin G. Ramayya, Isaac Pedisich, Michael J. Kahana *

Department of Psychology, University of Pennsylvania, Philadelphia, PA 19104, USA

ARTICLE INFO

Article history:
Received 24 October 2014
Accepted 19 April 2015
Available online 30 April 2015

Keywords:
Intracranial electroencephalography
ECoG
iEEG
High frequency activity
HFA
Reward
Value
Valence
Reinforcement learning

ABSTRACT

The brain's sensitivity to unexpected gains or losses plays an important role in our ability to learn new behaviors (Rescorla and Wagner, 1972; Sutton and Barto, 1990). Recent work suggests that gains and losses are ubiquitously encoded throughout the human brain (Vickery et al., 2011), however, the extent to which reward expectation modulates these valence representations is not known. To address this question, we analyzed recordings from 4306 intracranially implanted electrodes in 39 neurosurgical patients as they performed a two-alternative probability learning task. Using high-frequency activity (HFA, 70–200 Hz) as an indicator of local firing rates, we found that expectation modulated reward-related neural activity in widespread brain regions, including regions that receive sparse inputs from midbrain dopaminergic neurons. The strength of unexpected gain signals predicted subjects' abilities to encode stimulus–reward associations. Thus, neural signals that are functionally related to learning are widely distributed throughout the human brain.

Introduction

Theories of reinforcement learning postulate that greater learning occurs following unexpected outcomes than following expected outcomes (Rescorla and Wagner, 1972; Pearce and Hall, 1980; Sutton and Barto, 1990). How the brain represents these unexpected gains and losses has been the focus of considerable research. For example, functional neuroimaging studies have identified a specialized group of brain regions that encode reward prediction errors (Berns et al., 2001; McClure et al., 2003; Pessiglione et al., 2006; Montague et al., 2006; Rutledge et al., 2010; Bartra et al., 2013). Several of these regions (e.g., ventral striatum, medial prefrontal cortex) receive prominent inputs from midbrain dopaminergic (DA) neurons, a neural population known to be functionally important for reinforcement learning in animals (Schultz et al., 1997; Reynolds et al., 2001) and humans (Zaghloul et al., 2009; Ramayya et al., 2014a).

Recent evidence raises the possibility that the neural processes that support reinforcement learning may extend beyond regions that are heavily innervated by dopamine neurons. Vickery et al. (2011) used multi-voxel pattern analysis to decode outcome valence from activity in almost every cortical and subcortical region in the human brain. However, because this study did not assess reward expectation, the extent to which these widespread valence signals reflect reward prediction errors that are functionally important for learning is not known. If reinforcement learning is a widespread brain process, one would predict that valence representations throughout the brain would be modulated by reward expectation.

To test this hypothesis, we obtained intracranial electroencephalography (iEEG) recordings from the cortex and medial temporal lobe (MTL) of 39 patients with drug-refractory epilepsy as they performed a two-alternative probability learning task. We studied changes in high-frequency activity (HFA; 70–200 Hz) at individual electrodes, an established indicator of local spiking activity (Manning et al., 2009; Ray and Maunsell, 2011) that can be used to study heterogeneous patterns of activity within a region (Bouchard et al., 2013a). We identified putative valence signals that demonstrated differential HFA following positive and negative outcomes and we then assessed their relation to trial-by-trial estimates of reward expectation. In this way, we sought to characterize the anatomical distribution of expectation-modulated valence signals and assess their functional relevance for learning.

Materials and methods

Subjects

Patients with drug-refractory epilepsy underwent a surgical procedure in which grid, strip, and depth electrodes were implanted in order to localize epileptogenic regions. Clinical circumstances alone

* Corresponding author at: Department of Psychology, University of Pennsylvania, 3401 Walnut St., Room 303C, Philadelphia, PA 19104, USA.
E-mail address: kahana@psych.upenn.edu (M.J. Kahana).

http://dx.doi.org/10.1016/j.neuroimage.2015.04.037
1053-8119/

determined number of implanted electrodes and their placement. Data were collected from Thomas Jefferson University Hospital (TJUH) and the Hospital of University of Pennsylvania (HUP) in collaboration with the neurology and neurosurgery departments at each institution. Our research protocol was approved by the Institutional Review Board at each hospital and informed consent was obtained from the participants. In total, we recorded neural activity from 39 subjects (12 females, seven left-handed, mean age 37 years).

Reinforcement learning task

Subjects performed a two-alternative probability learning task, which has been previously used to study reinforcement learning and value-based decision making (Fig. 1; Frank et al., 2004, 2007; Zaghloul et al., 2012). During the task, subjects selected between pairs of Japanese characters ("items") and received positive or negative feedback following each choice. Subjects were informed that one item in each pair carried a higher probability of positive feedback than the other item, and were asked to select items that maximized their probability of obtaining positive feedback. On a given trial, the items were simultaneously displayed on the screen; one on the left side and one on the right side. They were presented on a dark gray background in white font. The items remained on the screen until subjects responded by pressing the left or right "SHIFT" button on a keyboard (to select the item on the left or right side of the screen, respectively). Once a response was registered by the computer, the selected item was highlighted in blue, and feedback was provided immediately. In the event of positive feedback, we presented a green screen and the sound of a cash register. In the event of negative feedback, we presented a red screen and the sound of an error tone. The colored screen was presented for 2 s. There was a 0–400 ms jitter between successive trials. Items were randomly arranged on the left or right side of the screen from trial to trial.

During a session, subjects were presented with up to three novel item pairs, each carrying a distinct relative reward rate (80/20, 70/30, or 60/40). This feature of the task allows for the study of value-based decision making in a subsequent stage of the experiment that is not considered in this study (Frank et al., 2007; Zaghloul et al., 2012). Distinct item pairs were presented in a randomly interleaved manner. Reward rates associated with each item were determined randomly prior to

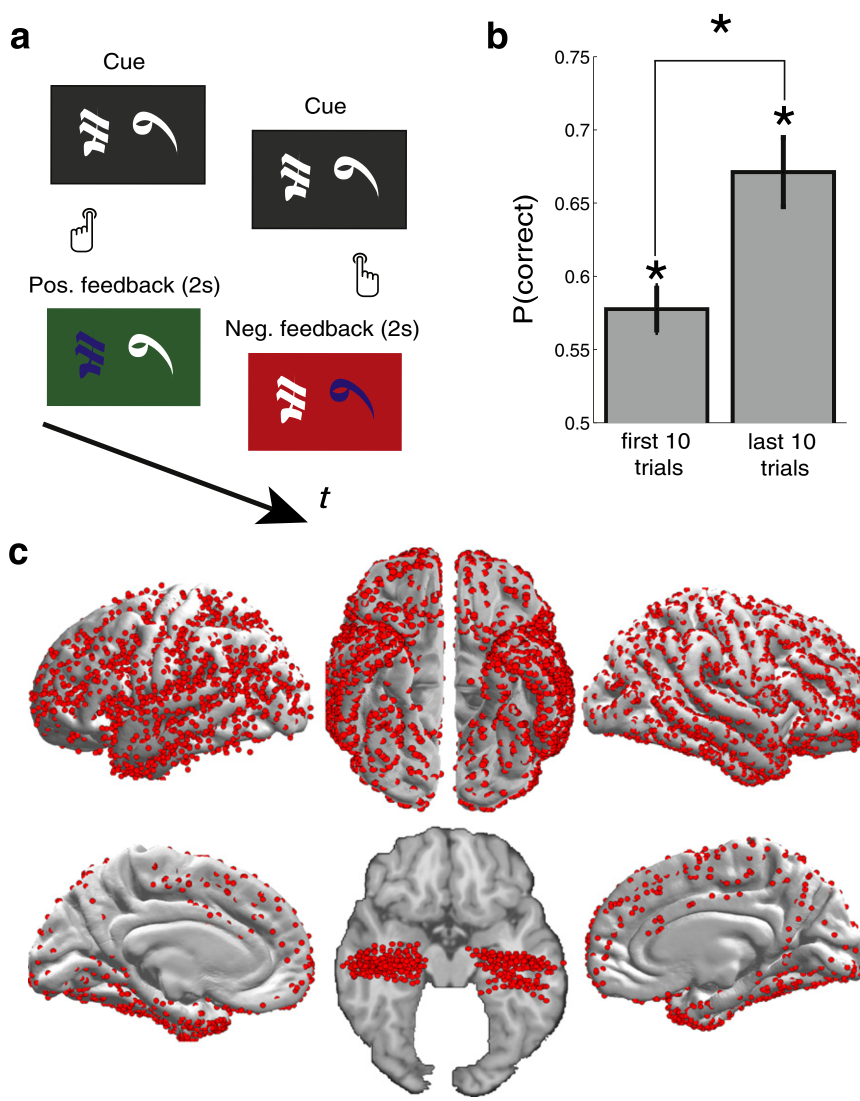

Fig. 1. Reinforcement learning task, and subjects' behavior, and electrode locations. a. Subjects selected between pairs of Japanese characters on a computer screen and probabilistically received positive or negative audio-visual feedback following each choice. b. Average tendency towards selecting the high-probability item during the first and last 10 trials of each item pair. Error bars represent s.e.m. across subjects. c. iEEG electrodes from each subject were localized to a common anatomical space (see Materials and methods section). We show strip and grid electrodes on the cortical surface, and depth electrodes targeting the medial temporal lobe on the axial slice. On rare occasions, depth electrodes were placed in the frontal and parietal lobes to supplement surface recordings (not shown).