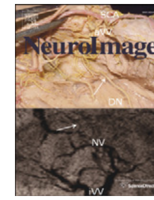
NeuroImage 70 (2013) 66–79

Contents lists available at SciVerse ScienceDirect

NeuroImage

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



Action selection in multi-effector decision making

Seth Madlon-Kay [a,*], Bijan Pesaran [b], Nathaniel D. Daw [b,c]

[a] Center for Cognitive Neuroscience, Duke University, Durham, NC 27708, USA
[b] Center for Neural Science, New York University, New York, NY 10003, USA
[c] Department of Psychology, New York University, New York, NY 10003, USA

ARTICLE INFO

Article history:
Accepted 2 December 2012
Available online 7 December 2012

Keywords:
Bimanual
Decision making
Value maps
Effector specificity
vmPFC

ABSTRACT

Decision making and reinforcement learning over movements suffer from the curse of dimensionality: the space of possible movements is too vast to search or even represent in its entirety. When actions involve only a single effector, this problem can be ameliorated by considering that effector separately; accordingly, the brain's sensorimotor systems can subdivide choice by representing values and actions separately for each effector. However, for many actions, such as playing the piano, the value of an action by an effector (e.g., a hand) depends inseparably on the actions of other effectors. By definition, the values of such coordinated multi-effector actions cannot be represented by effector-specific action values, such as those that have been most extensively investigated in parietal and premotor regions. For such actions, one possible solution is to choose according to more abstract valuations over different goods or options, which can then be mapped onto the necessary motor actions. Such an approach separates the learning and decision problem, which will often be lower-dimensional than the space of possible movements, from the multi-effector movement planning problem. The ventromedial prefrontal cortex (vmPFC) is thought to contain goods-based value signals, so we hypothesized that this region might preferentially drive multi-effector action selection.
To examine how the brain solves this problem, we used fMRI to compare patterns of BOLD activity in humans during reward learning tasks in which options were selected through either unimanual or bimanual actions, and in which the response requirements in the latter condition inseparably coupled valuation across both hands. We found value signals in the bilateral medial motor cortex and vmPFC, and consistent with previous studies, the medial motor value signals contained contra-lateral biases indicating effector-specificity, while the vmPFC value signals did not exhibit detectable effector specificity. Although neither region's value signaling differed significantly between bimanual and unimanual conditions, the vmPFC value region showed greater connectivity with the medial motor cortex during bimanual than during unimanual choices. The specific region implicated, the anterior mid-cingulate cortex, is thought to act as a hub that links subjective value signals to motor control centers. These results are consistent with the idea that while valuation for unilateral actions may be subserved by an effector-specific network, complex multi-effector actions preferentially implicate communication between motor regions and prefrontal regions, which may reflect increased top-down input into motor regions to guide action selection.

Introduction

Humans and other animals possess a variety of effectors, which together support a wide array of possible motor actions. The large space of possible movements implies a classic “curse of dimensionality” due to the exponential explosion of combinations of movements. For instance, coarsely considering only actions involving the eyes, arms, hands and legs yields a seven-dimensional space that can contain many millions of possible joint actions, even if each effector considered alone has only a few candidate movements. Given that the brain most certainly parses the body more finely than merely seven effectors, the actual number of possibilities must be many orders of magnitude larger. Representing such a large number of actions, let alone deciding between them, is computationally challenging. The problem becomes especially laborious when appropriate decisions are learned by trial and error, since even trying out all the possibilities is clearly impossible.

One method by which the brain seems to cope with this curse is by subdividing decisions into lower dimensional, effector-specific subspaces: that is, controlling particular effectors separately and independently from the others. Indeed, the basic organization of the brain seems well suited to a “divide and conquer” strategy. Primary motor cortex and many affiliated territories including basal ganglia, supplementary and premotor areas, contain effector-specific topographies wherein different muscle groups in different parts of the body are governed by spatially segregated populations of neurons (Alexander et al., 1986; Blanke et al., 2000; Leyton and Sherrington, 1917; Pesaran et al., 2006; Schlag

* Corresponding author at: 6 Washington Place rm 809, New York, NY 10003, USA.
E-mail addresses: seth.madlonkay@duke.edu (S. Madlon-Kay), bijan@nyu.edu (B. Pesaran), nathaniel.daw@nyu.edu (N.D. Daw).

1053-8119/$ – see front matter © 2012 Elsevier Inc. All rights reserved.
http://dx.doi.org/10.1016/j.neuroimage.2012.12.001

and Schlag-Rey, 1987). Similarly effector-specific systems can be found in sensorimotor areas such as posterior parietal cortex, where saccadic eye movements and reaches are represented by distinct subregions (Batista et al., 1999; Calton et al., 2002; Levy et al., 2007; Snyder et al., 1997). Recent evidence suggests that such effector-specific movement representations are also a substrate for evaluating and choosing among these actions, e.g. with the eye movement representations in parietal cortex carrying 'value maps' of rewards expected for different saccades and with action values for movements of the hands and eyes appearing in distinct subregions of supplementary motor cortex (Andersen and Buneo, 2002; Palminteri et al., 2009; Platt and Glimcher, 1999; Scherberger and Andersen, 2007; Shadlen and Newsome, 2001; Sugrue and Corrado, 2004; Wunderlich et al., 2009). Furthermore, in a recent fMRI study we tested explicitly for subdivision across effectors (here, left and right hands) and found lateralized patterns of BOLD signals related to valuation and reward prediction errors consistent with the separation of a multi-effector task into effector-specific subproblems (Gershman et al., 2009).

However, separating the control of effectors in this way introduces a related problem of coordination. Many actions require the conjoint action of multiple effectors working together to accomplish a single goal. When playing piano chords, for example, playing a C with the left hand will sound dissonant if the right hand chooses a B but pleasingly consonant if instead the right hand plays E. In these cases, the value of a particular effector's movements depends on the action of another effector. This means that the fundamental units of evaluation and decision must be combinations of movements spanning multiple effectors, and cannot be represented by separate maps of each effector's outcomes alone. The need for such coordination is also illustrated by corpus callosotomy patients. While they often show improved performance over neurologically intact controls in tasks in which the two hands must act completely independently, such as finger-tapping with different rhythms for each hand, they are unable to perform or learn novel actions requiring that the two hands act in concert, such as threading a needle or playing the piano (Brinkman, 1984; Franz et al., 2000; Kennerley et al., 2002). In other words, for the case of novel bimanual actions, such patients appear limited to actions that can be represented by independent, effector-specific systems.

How, then, does the brain evaluate high dimensional multi-effector actions? One approach is simply also to maintain, in the motor system, representations over multi-effector actions, and their values, as well. Indeed, single units in areas such as the supplementary motor area have been reported with differential tuning for bimanual movements relative to their unimanual counterparts (Donchin et al., 2002; Tanji et al., 1988). Such a solution would be most useful for selecting among sets of commonly-used or well-learned multi-effector actions, or within groups of effectors commonly used together. However, if applied to all possible multi-effector movements, this solution would likely run into of the curse of dimensionality because the space of all possible multi-effector actions is so vast.

Another approach, which may be complementary, is to rely on a "goods-based" rather than an "action-based" evaluation strategy (Padoa-Schioppa and Assad, 2006; Wunderlich et al., 2010): representing the value of a few possible outcomes or goods that might serve as the goals of actions, separate from the movements needed to obtain them. Once a choice has been made in the space of outcomes (which may be low-dimensional), the motor system is left to solve the more targeted problem of executing the choice. Intuitively, this is like first choosing between peanut butter and potato chips, and subsequently planning the movements required to grab the chosen snack food and extract it from its container. Such an approach would support a different sort of divide and conquer strategy, whereby learning, evaluation and choice over goals could be separated from the movement programming required to obtain them. The ventromedial prefrontal and orbitofrontal cortices (vmPFC and OFC) are thought to contain this sort of "goods-based" value signal. Neurons in the primate OFC represent values of particular consumption goods regardless of the actions involved in obtaining them, and their firing properties generally do not vary with the motor contingencies of tasks (Kennerley and Wallis, 2009; Padoa-Schioppa and Assad, 2006; Wallis and Miller, 2003). Similarly, in the ventromedial prefrontal cortex in humans, BOLD activity correlates with subjective value across a very wide and general variety of contexts, including those that do not involve making a motor action (Berns et al., 2001; Kable and Glimcher, 2007; Knutson et al., 2000; Levy and Glimcher, 2011; O'Doherty et al., 2002; Peters and Büchel, 2009; Plassmann et al., 2007; Tom et al., 2007; Wunderlich et al., 2010). The distinction between goods- and action-based choice models has been hotly debated in decision neuroscience (Glimcher, 2008; Padoa-Schioppa and Assad, 2006; Platt and Padoa-Schioppa, 2008), but these mechanisms need not be mutually exclusive, and recent imaging studies suggest that the brain uses both (Wunderlich et al., 2009, 2010). In particular, for the reasons discussed above, these two mechanisms are well suited to single- and multi-effector problems, respectively. We thus hypothesized that the vmPFC might preferentially guide action selection during multi-effector choice.

To investigate the neural mechanisms of multi-effector decision making, we used fMRI to examine patterns of neural activity in humans playing a four-armed bandit reinforcement learning task, in which choices were executed, in different conditions, by unimanual or bimanual button presses. Although in both conditions subjects learned the values of four shapes, in the bimanual condition, the response requirements for choosing a shape were such that (unlike the separable bimanual task studied by Gershman et al. (2009)) the value of each action was indivisible across the effectors. For this reason, choice could operate over effector-specific value representations in the unimanual condition, but not in the bimanual one. (Note that here we are conceptualizing each hand as a single effector; however, one could instead describe each finger as an effector and the above claims would still apply.) Using estimates of the participants' values of the options on each trial, we probed the brain for differences in patterns of value-related activity associated with bimanual versus unimanual actions. In particular, we examined functional connectivity between value-related areas of vmPFC and more posterior motor regions as a potential index of the involvement in action of abstract valuation systems.

Methods

Participants

Twenty-three right-handed adults, ages 19 to 38 (median 22, 17 females), were recruited from the New York University community to participate in the experiment in exchange for payment. Participants received a fixed amount of money for completing the experiment as well as the money they won during the task. Three participants were excluded due to technical difficulties, leaving twenty subjects whose data were used in the analysis. All participants gave informed consent, and the study was approved by the New York University Committee on Activities Involving Human Subjects.

Experimental task

Participants performed a bandit task in which they earned money based on their choices. On each trial, participants selected one of four shapes, each of which was associated with some probability of winning money (Fig. 1A). Participants received feedback regarding whether they won money on that trial 4 s after the choice was executed. The probability of winning from each shape changed slowly and independently over the course of the task according to a Gaussian random walk, with reflecting bounds below at 0.3 and above at 0.7, so maximizing payoff required continuously tracking shape values throughout the task. For each winning trial participants received \$0.10 at the end of the experiment.