Cognition 129 (2013) 563–568

Contents lists available at ScienceDirect

Cognition

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

Brief article

# The influence of depression symptoms on exploratory decision-making

Nathaniel J. Blanco [a], A. Ross Otto [b], W. Todd Maddox [c,*], Christopher G. Beevers [c], Bradley C. Love [d,*]

[a] Department of Psychology, University of Texas at Austin, United States
[b] Center for Neural Science, New York University, United States
[c] Department of Psychology, Institute for Mental Health Research, University of Texas at Austin, United States
[d] Department of Cognitive, Perceptual and Brain Sciences, University College London, UK

ARTICLE INFO

*Article history:*
Received 2 March 2013
Revised 13 August 2013
Accepted 14 August 2013
Available online 17 September 2013

*Keywords:*
Exploration vs. exploitation
Depression
Reinforcement learning
Dynamic decision making
Cognitive modeling

ABSTRACT

People with symptoms of depression show impairments in decision-making. One explanation is that they have difficulty maintaining rich representations of the task environment. We test this hypothesis in the context of exploratory choice. We analyze depressive and non-depressive participants' exploration strategies by comparing their choices to two computational models: (1) an "Ideal Actor" model that reflectively updates beliefs and plans ahead, employing a rich representation of the environment and (2) a "Naïve Reinforcement Learning" (RL) model that updates beliefs reflexively utilizing a minimal task representation. Relative to non-depressive participants, we find that depressive participants' choices are better described by the simple RL model. Further, depressive participants were more exploratory than non-depressives in their decision-making. Depressive symptoms appear to influence basic mechanisms supporting choice behavior by reducing use of rich task representations and hindering performance during exploratory decision-making.



## 1. Introduction

Depression is a common condition linked to suicide attempts, interpersonal problems, unemployment, and substance abuse (Kessler et al., 2003). The World Health Organization estimates that 121 million people suffer from depression and many more have elevated depressive symptoms. Clarifying the relationship between depressive symptoms and cognition may be useful in understanding both depression and basic cognitive processes.

Depressive symptoms are associated with lower performance in working memory (Rogers et al., 2004), problem-solving (Elderkin-Thompson, Mintz, Haroon, Lavretsky, & Kumar, 2006), planning (Elliott et al., 1997), and decision-making tasks (Clark, Chamberlain, & Sahakian, 2009; Gradin et al., 2011; Maddox, Gorlick, Worthy, & Beevers, 2012; Murphy et al., 2001; Pizzagalli, Iosifescu, Hallett, Ratner, & Fava, 2008). Computational models have proven useful in understanding the basis for these impairments, particularly in decision-making (Eshel & Roiser, 2010; Montague, Dolan, Friston, & Dayan, 2012). Along these lines, Paulus and Yu (2012) suggest that depressive symptoms alter action-value computations, causing abnormal decision-making in people with higher depressive symptoms.

Here we examine how these computations in exploratory choice are affected by depressive symptoms. In decision-making, optimal choice often requires building a representation of the task that supports effective planning. Because depressive individuals exhibit deficits in planning and working memory, we expect that they will have difficulty maintaining a rich representation of the

* Corresponding authors. Address: University College London, Department of Cognitive, Perceptual and Brain Sciences, 26 Bedford Way, London WC1H 0AP, UK. Tel.: +44 020 7679 2000 (B.C. Love).

*E-mail addresses:* Maddox@psy.utexas.edu (W.T. Maddox), b.love@ucl.ac.uk (B.C. Love).

0010-0277/$ - see front matter Crown Copyright © 2013 Published by Elsevier B.V. All rights reserved.
http://dx.doi.org/10.1016/j.cognition.2013.08.018

task environment (Otto, Gershman, Markman, & Daw, 2013). If so, they should tend to rely on simple strategies and impoverished task representations, resulting in suboptimal decision-making (Montague et al., 2012).

Supporting this notion, Huys et al. (2012) found that in sequential decision-making, depressive symptom severity correlated with a tendency to "prune" (i.e. avoid mentally searching) paths that included a large loss even when it was advantageous to consider such options. Huys et al. suggested that pruning is an inflexible strategy not adaptive to tasks demands that is *reflexively* applied in response to punishment. The greater the depressive symptoms, the stronger the tendency was to prune, as opposed to *reflectively* consider alternative strategies and plan.

Montague et al. (2012) further suggest depressive decision-makers should explore less. Indeed, people with depressive symptoms exhibit less switching between options in a choice task where reward contingencies change over time (Cella, Dymond, & Cooper, 2010). Moreover, the increased pruning (which reduces the number of solutions considered) by depressives in Huys et al. is akin to reduced exploration, though these participants sample more uniformly (i.e., are more exploratory) within this reduced choice set.

Here, we provide a finer-grained examination of exploration strategies in depressives' decision-making, directly testing the hypothesis that individuals with depressive symptoms are less likely to use rich task representations to reflectively update their beliefs and plan choices. Our computational approach affords understanding the basis of deficits in those suffering from depressive symptoms and the nature of exploratory choice more generally.

### 1.1. The Leapfrog task

We examine the effects of depressive symptoms on exploratory strategies by using a paradigm termed the "Leapfrog" task (Knox, Otto, Stone, & Love, 2012), a variant of the "bandit" task (Sutton & Barto, 1998). In this task (Fig. 1), one of two options gives a higher reward than the other. On any trial the currently inferior option can increase in value, becoming the better option. This change happens with a fixed probability called the "volatility" of the environment. Because the relative superiority of the options shifts over time, on each trial the participant must choose between *exploiting* the option with the highest observed reward and *exploring* to see if the other option has surpassed it. Because each choice is effectively reduced to the decision to explore or exploit, the Leapfrog task is well-suited to investigating exploratory behavior.

### 1.2. Reflexive vs. reflective strategies

Recent work has sought to characterize the types of behavior and/or representations that give rise to exploratory choice (Badre, Doll, Long, & Frank, 2012; Otto, Markman, Gureckis, & Love, 2010). One basic theoretical distinction is whether exploration is guided by beliefs that evolve in a principled manner to reflect uncertainty in the environment (i.e., *reflective* updating) or by beliefs that only change as a result of direct feedback (i.e., *reflexive* updating; Knox et al., 2012). The reflective and reflexive conceptualization echoes the distinction between "model-based" and "model-free" learning in Reinforcement Learning (RL; Daw, Niv, & Dayan, 2005).

A *reflexive* learner has no representation of its environment besides expected values for each option, which are updated only after receiving rewards. State uncertainty is not utilized to guide decisions. Exploratory choices are undirected, resulting from a purely stochastic decision process. In *reflective* choice, conversely, the learner has a richer representation of their environment. This representation can include (among other things) beliefs about the state of the environment, state transitions, and the probabilities of events. In the Leapfrog task, a reflective learner could direct its choices according to its belief as to whether the observed-to-be-superior option is still superior. With each successive exploitive choice, the probability that the relative superiority of the options has flipped increases, making the state of the environment less certain. In this way, exploratory behavior can be directed by uncertainty; as uncertainty increases, exploration becomes more valuable. By using this knowledge about the environment to plan exploration, reflective strategies should outperform reflexive strategies on the Leapfrog task.

Through quantitatively comparing how well a reflexive vs. reflective account characterizes an individual's choices, we assess whether the relative usage of reflexive and reflective strategies differs between depressive and non-depressive individuals. As previously discussed, we predict that depressive individuals will be less likely to use reflective strategies.

### 1.3. Models evaluated

We fit computational models that embody reflective and reflexive strategies to participants' data to evaluate their strategies. The "Ideal Actor" reflectively updates beliefs and plans ahead, taking into account the information gained by each choice and making choices that maximize long-term payoffs. Action-values are a product of both expected rewards and the potential to reduce uncertainty about the state of the environment. In contrast, the Naïve RL model instantiates the reflexive account of choice, in which the values of actions are based only on the rewards experienced so far. Its beliefs are updated reflexively in response to observed changes in rewards.

Turning to the model details, both models incorporate a Softmax choice rule (Sutton & Barto, 1998), which chooses options as a function of the computed action-values. Critically, the action-values used in the Softmax choice rule differ between the two models, leading to qualitative differences in exploratory behavior. The Naïve RL model explores with equal probability on every trial. For the Ideal Actor model, the probability of exploring increases after each successive exploitive choice (see Fig. 4A).

For the Naïve RL model the value of each action is equal to the last observed reward for that action. Algorithmically, it is equivalent to the Softmax model used in Worthy, Maddox, and Markman (2007) with a learning rate of 1. The Ideal Actor computes action-values in two steps. First, it optimally updates its (Bayesian) beliefs about the state of