Elevated activation of dopaminergic brain areas facilitates behavioral state transition

Beata J. Grzyb*†, Joschka Boedecker†, Minoru Asada†, and Angel P. del Pobil*
*Robotic Intelligence Laboratory, Jaume I University
Castellon de la Plana, 12071, Spain
Email: {grzyb,pobil}@icc.uji.es
† JST ERATO Asada Project, Graduate School of Eng., Osaka University
2-1 Yamadaoka, Suita, Osaka, 565-0871 Japan
Email: {joschka.boedecker,asada}@ams.eng.osaka-u.ac.jp

***Abstract*—Dopamine neurons appear to code the discrepancy between the reward and its prediction, and as such play a key role in learning from positive and negative feedback. Although the traditional view stresses the role of errors in learning, we suggest that a temporal decrease in learning from negative feedback may in fact facilitate the process of finding more suitable behaviors that would reflect the change in behavioral competences of the agent. The major premise of our approach is that omission of the errors enables selection of different behaviors in a context when they normally would not be selected. Therefore, it provides more learning opportunities for fine-tuning these behaviors. We propose that omission of the errors is tightly related to an elevated level of dopamine that is caused by a high reward for gaining the control over the enviromnent. Our results with a robot simulator serve as a proof-of-concept for our approach.**

I. Introduction

Obtaining a desired objective is a satisfying experience for nearly everyone. The way of achieving the goal is not always easy and straightforward. Sometimes many different methods need to be explored until the goal can be acomplished. If none of the possible solutions is helpful in attaining the goal, the goal is memorized as impossible and later attempts abandoned. Sometimes, however, these efforts may be reassumed when new competences appear. This can clearly be seen in infants. At the beginning nothing is reachable for them, as they do not have enough skills to coordinate eye and hand movements. Soon these first movements become successful, and infants learn that only close but not far objects are reachable for them. This, however, changes with the onset of crawling and later on with walking behavior. Consequently, infants need to relearn that attaining distant objects is possible only by locomotion, a behavior more appropriate in this context.

Our previous experiments revealed that infants during the transition phase to walking show a decreased ability to learn what lies within their reachable space [1], [2]. We suggested that the blocked ability to learn from negative outcome while reaching makes infants fine-tune their walking skill, as a primary motive for walking is to reach for something [3]. Thus, we proposed that a temporal decrease in learning from negative feedback could be an efficient mechanism behind infant learning new skills. Furthermore, we proposed that disregarding the errors is tightly connected to the sense of control, and results from an extremely high level of self-efficacy. In this paper, we propose possible brain mechanisms that may lead to omission of errors during feedback processing.

An important theoretical framework underlying our proposal is the dynamic systems approach [4]. In this framework stable configurations of a system are referred to as attractors. In that sense, the decision not to reach for far objects is a stable attractor for the infants that are not able to locomote on their own. When a possibility of a new behavioral competence appears, for example walking, the system should change its decisions about appropriete behaviors. Therefore, the existing stable attractor needs to be destabilized in order for the system to accomodate a new one.

Another important role in our proposal is played by the neuromodulator dopamine. Phasic dopamine signals were suggested to trigger a switch in the current attractor state in the networks of prefrontal cortex, by transiently enhancing afferent input while potentiating local inhibitory signals thus gating new information into the prefrontal cortex [5]. Furthermore, dopamine in the amygdala was suggested to modulate behavioral transitions [6] that characterize development and progression in competences. The main focus in the literature, however, has been put on dopamine and its role in the prediction of rewards and determination of whether predictions about outcomes are violated or verified. A decrease of dopaminergic activity owing to the omission of rewards, was mainly interpreted as coding a prediction error or learning signal that is supposed to trigger learning and adaptation of future behavior [7]. Another interesting point is that the dopamine system may act at several different timescales in the brain from the fast, restricted signalling of reward and some attention-inducing stimuli to the slower processing of a range of positive and negative motivational events [8].

The main premise of our approach is that when an infant gains control over its environment, the reward circuitry in her brain will deliver a large reward to the executive brain areas facilitating repeated selection of actions that led to the gain of control. This reward is experienced even though behavior on a shorter time scale (e.g. reaching) fails, but progress is made on behavior spanning an extended time scale (e.g. reaching by walking). Omitting the errors in this way

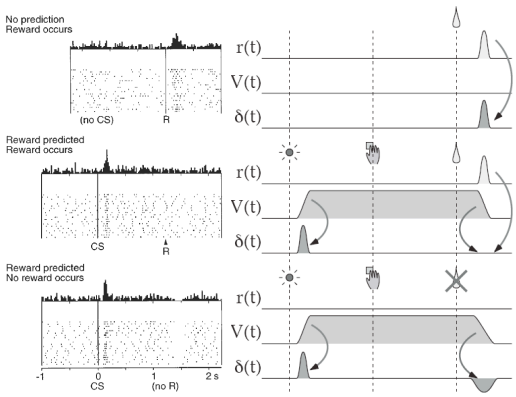


Fig. 1. On the left: Reward prediction error response of single dopamine neuron (taken from [9]). On the right: Interpretation of the responses of midbrain dopamine neurons in the TD model; r(t): reward; V(t): reward prediction; δ: dopamine response (taken from [10])

enables selection of different behaviors in a context when they normally would not be selected, thus destabilizing existing attractors and facilitating the formation of new ones.

We review the mechanism of the reward prediction error both in neuroscience, and in reinforcement learning theory in the next section. After that, we present cases where learning from negative outcome was signicantly decreased. In Sec. IV we introduce our hypothesized neural circuitry facilitating behavioral state transition. The next section, presents results from our robot simulation study where the underlying assumption of the model were tested. We close the paper with discussion and an outline of future work.

II. Prediction error learning

The response of dopamine neurons appears to code the discrepancy between the reward and its prediction [9]. A typical response of a single dopamine neuron is shown in Fig. 1 (on the left). During the acquisition process the dopamine neurons increase firing rates when reward (R) is received but not expected (no CS). Over time this increase in firing rate is back propagated to the earliest reliable stimulus (CS) for the reward. The dopamine cells no longer increase their firing rate upon presentation of the predicted reward. However, when rewards are expected but not received, the firing of dopamine neurons drops below tonic baseline levels.

The activity pattern of dopamine neurons represents the reward prediction error, which is central to the temporal difference (TD) learning model [11], [12]. The TD model calculates a prediction error $\delta(t)$ based on the temporal difference between the current discounted value function $\gamma V(t)$ and that of the previous time step, $V(t-1)$.

$$\delta_t = r(t) + \gamma V(t) - V(t-1), \tag{1}$$

where γ is a discount factor which allows rewards that arrive sooner to have a greater influence over delayed ones, and $r(t)$ represents the current reward [12], [10]. The interpretation of the dopamine neurons responses in the TD model is shown in Fig. 1 (on the right). Before learning, no reward is predicted, that is $V(t) \equiv 0$. Thus, the TD error $\delta(t)$ is the same as the reward itself. After learning has been completed, the predicted future reward $V(t)$ builds up immediately after the cue signal, causing the discounted temporal derivative to provide a positive pulse in the TD error even if there is no reward. At the time of reward delivery, $V(t)$ drops to zero and the negative temporal derivative of $V(t)$ cancels out the positive reward signal. However, when the reward is omitted, there is a negative response due to the drop in the predicted reward $V(t)$. By acting as a teaching signal, dopamine-mediated prediction errors are expected to gradually train learning mechanisms to improve their predictions in an incremental and trial-by-trial fashion [13].

III. Decreased ability of learning from errors

Although there may be some individual differences due to genetic variations affecting dopamine function, in general healthy people are equally good at learning to obtain positive outcomes and to avoid negative outcomes. People with Parkinson's disease, however, show specific deficits in trial-and-error learning from feedback. These effects were nicely explained by Frank's basal ganglia model [14]. Basal ganglia dopamine levels in these patients are severly depleted as a result of cell death. As the positive outcomes are signaled by a raise in the firing rate of dopamine neurons, the depleted overall dopamine levels in unmedicated patients results in a weaker reinforcement of the stimulus. On the other hand, the errors in reward prediction are signaled by a decrease in the firing rate of dopamine neurons. As a result of low dopamine levels, the errors in unmedicated patients have much stronger negative reinforcement of the stimulus. The dopaminergic medications, however, reverse these biases and medicated individuals with Parkinson's disease are better at learning from positive than from negative feedback. The dips of dopamine required to learn negative prediction errors are effectively filled in by the medication, and such blunting of negative prediction errors reduces learning from negative outcomes. Essentially, the medication prevents the brain from naturally and dynamically regulating its own dopamine levels, which has a detrimental effect on learning, particularly when dopamine levels should be low, as for negative decision outcomes.

The inability to learn from negative feedback was shown in healthy subjects during the trust game [15]. In this experiment information about the moral profile of the oponent was provided to the players before the game started. This information can create a prior belief, but feedback from the game should adjust this prior belief to reflect new evidence. However, the experiment showed the lack of differential responses between the positive and negative outcomes when playing with morally good or bad partners. More specifically the activation of the caudate nucleus differentiated between positive and negative feedback, but only for the 'neutral partner', and not for the 'good' one, and only weakly for the 'bad' one. The normal trial-and-error learning would predict a sharp decrease in the feedback response following violations of expectations. One of the possible explanations suggested by the authors was that

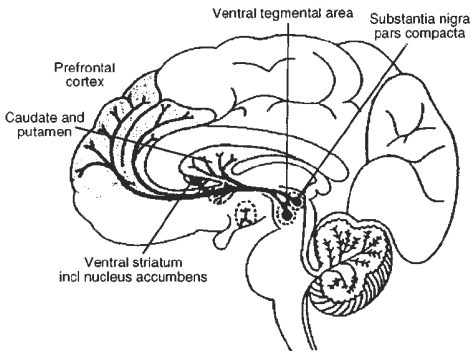


Fig. 2. The neurocircuitry of a reward system (taken from [21]).

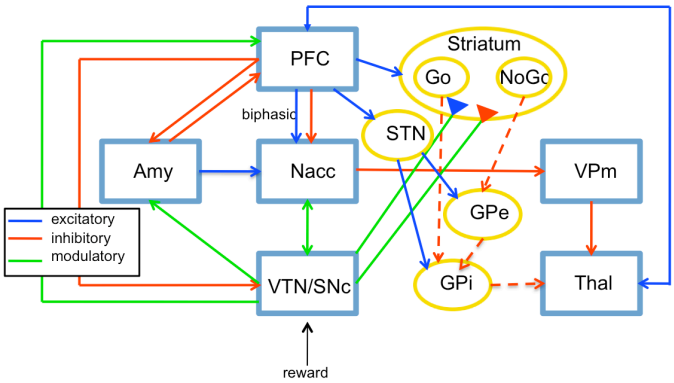


Fig. 3. The striato-cortical loops, including the direct ("Go") and indirect ("NoGo") pathways of the basal ganglia, and neural circuitry for perceiving control. PFC: prefrontal cortex; Amy: amygdala; Nacc: nuccleus accumbens; VTA: ventral tegmental area; SNc: substantia nigra pars compacta; GPe: internal segment of globus pallidus; GPe: external segment of globus pallidus; STN: subthalamic nucleus; Thal: thalamus; VPm: ventral pallidum.

participants had a reward reaction to the presentation of the morally good partner, irrespective of decision.

In patients with bipolar disorder, failures in motor learning may result from the lack of striatal error signal during unsuccessful motor inhibition. Such deficits in motor regulation could be related to the emotional disregulation, as irritability and decreased motor inhibition may be linked mechanistically [16]. The impulsivity was suggested to represent a core characteristic of the disorder and to be responsible for symptoms like hyperactivation, excitability, and hasty decision making [17]. Patients with bipolar mania tend toward high goal setting, have unrealistically high success expectancies [18], and exhibit increased goal-directed activity and excessive involvement in pleasurable activities that have a high potential of risk [19]. Bipolar patients show elevated activation of dopaminergic brain areas when expecting high rewards compared to anticipation of no rewards, which could result from dysfunctional nucleus accumbens activation during prediction error processing [20]. When both, schizophrenia patients and healthy controls, showed lower nucleus accumbens activation upon omission rather than upon receipt of rewards as a potential correlate of such a learning signal, bipolar manic patients did not display a similar reduction in the activation of dopaminergic brain regions.

We have presented different cases where learning from negative outcome was significantly decreased. The first lesson from these examples is that elevated state of dopaminergic areas can lead to omission of the errors during learning like in the case of Parkinson's patients. The second lesson is that, abnormal activity in the striatum (dorsal or ventral) also causes decreased ability to learn from negative feedback. We believe that temporal omission of errors while learning a new skill may result from a similar mechanism. The next section introduces the details of our hypothesis.

IV. Sense of Control and Omitting the Errors

The principal assumption behind our approach is that a need for control is innate, and exercising control is extremely rewarding and beneficial for an individual's wellbeing [22], and people's ability to gain and maintain a sense of control is essential for their evolutionary survival [23]. The hypothesized neural circuitry that would explain the facilitation of behavioral state transition is depicted in Fig. 3.

Similary to the proposed BG models (eg. [24]), there are two BG pathways to selectively facilitate the execution of the most appropriate motor commands ("Go" pathway), while suppressing competing commands ("NoGo" pathway). The "Go" pathway depends on D1 receptors and supports learning from positive feedback, whereas the "NoGo" pathway depends on dopamine D2 receptors and supports learning from negative feedback. These two pathways compete with each other when the brain selects among possible actions, so that an adaptive action can be facilitated while at the same time competing actions are suppressed. More specifically, striatal "Go" neurons directly project to and inhibit the *internal* segment of the globus pallidus (GPi). The GPi in turn disinhibits the thalamus eventually facilitating the execution of the motor commands. Contrary, striatal "NoGo" neurons project to and inhibit the *external* segment of globus pallidus (GPe), releasing the inhibition of GPe onto GPi, and thus blocking the motor activity. Dopamine modulates the relative balance of these pathways by exciting synaptically-driven activity in Go cells via D1 receptors, while inhibiting NoGo activity via D2 receptors.

Prefrontal cortex (PFC) is constantly involved in the acquisition of new skills and knowledge, and may also play a role in organizing other parts of the cortex [25]. Increased activity in the medial PFC has been associated with perception of control [22]. The PFC and the amygdala have synergistic roles in regulating purposive behavior [26]. While the PFC guides a goal-directed behavior, the amygdala appears to extract the affective significance of stimuli. Communication between these two brain regions is bidirectional and appears to be essential in judging rewarding or aversive outcomes of actions. The PFC was shown to inversely correlate with amygdala during successful emotion regulation [27]. The inverse relationship reflects the inhibitory pathway from the dorsal and lateral regions of PFC to the amygdala. Furthermore, it

was proposed that amygdala drives vmPFC in a bottom-up affective reactivity task but can be downregulated by more dorsal and lateral portions of the PFC via the vmPFC in a top-down reappraisal task. The optimal balance between such bottom-up and top-down influences in a given emotional situation was suggested to be crucial for the individual to respond adaptively [28].

The nucleus accumbens (Nacc) is a hub for information related to reward, motivation, and decision making [29]. The Nacc provides a ventral pathway by which the limbic system and prefrontal areas can influence the initiation of goal-directed behavior [30]. Dopamine D1 and D2 agonist when injected in the Nacc compared to the dorsal striatum facilitate the initiation, speed and vigor of locomotion, and markedly increase the frequency and duration of spontaneous exploratory activity. Suppression of ventral striatal activity when anticipated rewards were not obtained has been interpreted as a prediction error signal [20]. The Nacc receives strong, direct projection from the amygdala and prefrontal cortex. The PFC modulation of Nacc dopamine function appears to be biphasic [31]. Under normal activity PFC provides an inhibitory control over Nacc dopamine release. Electrical stimulation of PFC at $10Hz$, which closely corresponds to the firing rate of PFC neurons in animals engaged in cognitive tasks decreases dopamine release in the NAcc. However, electrical stimulation at $60Hz$ that is much higher then normal activity, caused an increase in NAcc dopamine levels. Activated Nacc neurons project to and inhibit pallidal neurons in the region called ventral pallidum (VPm). The suppression of tonic activity in the pallidum then disinhibits the thalamic nucleus [32].

The ventral tegmental area (VTA) dopamine cells play a crucial role in facilitating motivated behavior via its coordinated modulation of prefrontal and Nacc circuity, as well as its direct input to limbic structures which effects input to the Nacc at source [33]. Moreover, with simultaneous stimulation of both the amygdala and VTA, Nacc stimulation more readily produces initiation of forward locomotion and exploratory activity to novelty [30]. Dopaminergic input from the VTA modulates the activity of neurons within the nucleus accumbens, as well as within the PFC [34].

One possible explanation for decreased learning from negative feedback is that exercising control is highly rewarding itself and even if the outcome of the action is not as predicted, still the reward for gaining control is provided. That leads to high activity in the PFC. As discussed previously, the PFC modulates the Nacc dopamine function. This regulation is biphasic, and at normal activity the PFC provides an inhibitory control over Nacc dopamine release, but the PFC stimulation at much higher than normal levels increases nucleus accumbens dopamine. Herein, we assume that gaining control evokes such a high PFC response. Thus, high activity in Nacc leads to disinhibition of the VPm, and in turn dishinhibition of the thalamus. Simply speaking, that facilitates selection of the behaviors that led to the gain in control. This loop bypasses the striatal areas involved in action selection (colored yellow in Fig. 3). However, the dopamine prediction error that helps to improve the selected behavior still reaches this areas. Our hypothesized role of ignoring the errors is important only in the more executive areas responsible for action selections.

Fig. 4. The M3-neony robot simulator.

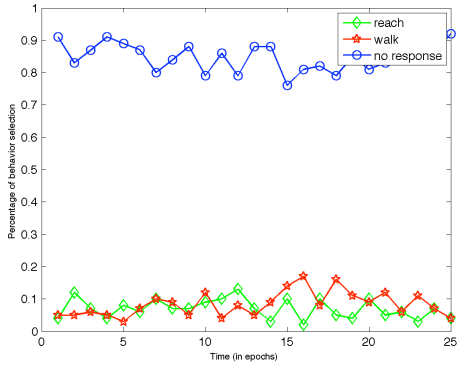
The details of this model are still to be verified, but its underlying assumptions about the role of ignoring the errors during hierarchical skill acquisition have been tested in a simulation study outlined in the next section.

V. Simulation

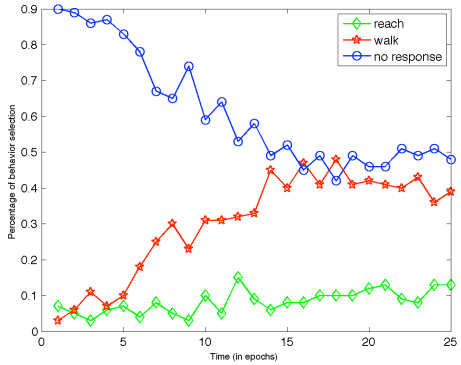
We investigated how ignoring the errors could help a robot (shown in Fig. 4) to learn new skills in an approximate optimal control framework. For the purpose of our study, the framework had a two-layer structure. The top layer, was a decision making layer, that was trained using standard Q-learning to select appropriately for a given context, one of the three possible behaviors, that is reaching, walking or no response. Herein, we made use of a standard inverse kinematics controller for the reaching action, and only the walking module was trained using standard Q-learning.

The state space of the decision making module was a discretized distance to the goal (6 states in our case changing by $2cm$). The goal of the modul was to select one of the possible sub-modules depending on their predicted action outcome. The module received a reward ($R = 60$) when the selected action was successful, and a punishment ($R = -30$) in the opposite case. The walking module had 6 different predefined states and actions, each state was described by 8 joint angles (4 for each leg). The goal of the module was to learn how to alternate from one state to another so that the robot does not loose balance, and it moves forward at the same time. The module received a partial reward for getting closer to the goal ($r = 10$), and negative reward for moving backwards ($r = -3$). When the robot reached the goal the module received additional reward ($r = 60$). Any action that resulted in loosing balance was punished ($r = -30$). In the simulations, epsilon greedy action selection was used with $\epsilon = 0.1$.

The simulation started with a robot not able to walk. The action of walking was available for selection, but its execution resulted in no movement. We simulated the onset of walking at $w = 40$ epochs. Until the onset of walking the distance to the

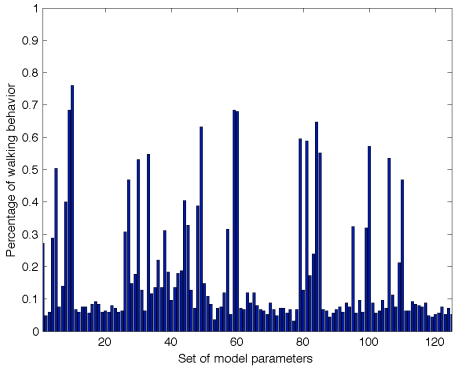


(a) The robot without the state of elation.

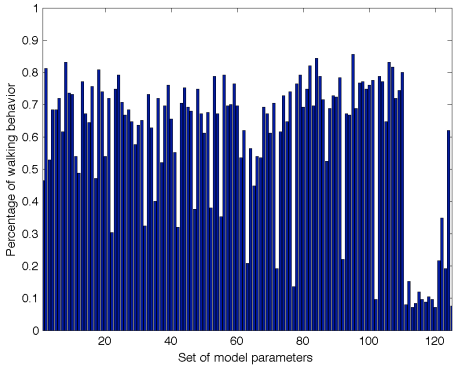


(b) The robot with the state of elation.

Fig. 5. The percentage of behavior selection.



(a) The robot without the state of elation.



(b) The robot with the state of elation.

Fig. 6. The percentage of walking behavior selection.

object (close or far distance) was changed randomly with 40% probability of change. After the onset of walking the object was placed only far away from the robot. We tested the robot in two different scenarios: without the state of elation, and with the state of elation. The state of elation was simulated by ignoring the negative outcomes of the actions in the decision making layer.

The settings and thus the behavior of the robot before the onset of walking was the same in both scenarios. Therefore, only the results of the simulations after this period are shown. As the robot chose actions with certain probability, the results of the simulations may vary across trials. We present the average results over 10 different trials. As it can easily be seen in Fig. 5(a), the robot without elation learned that the object is not reachable, and the probability of selecting the "no response" behavior was very high during the entire experiment. The robot had almost no opportunities to practice the walking behavior. On the other hand, the robot with the state of elation (shown in Fig. 5(b)), after 13 epochs started to select walking behavior more frequently making it possible for the walking module to improve.

As the results of the simulations may strongly depend on the values of reward, we repeated the simulation for different configurations of rewards. We varied the values for partial reward for getting closer to the goal in a walking module ($r \in \{5, 10, 15, 20, 25\}$), and the reward for successful action selection ($R \in \{30, 60, 90, 120, 240\}$) and punishment ($P \in \{-30, -60, -90, -120, -240\}$) for failure in reaching a goal in the decision making module. The results for total of125 different configurations are shown in Fig. 6(a) and Fig. 6(b). As it can be seen, just a few configurations for the robot without elation allow the walking behavior to be selected more often. Thus, introducing the state of elation, facilitated in many cases the transition from selecting no response behavior towards selection of the walking behavior.

VI. Discussion

In terms of the dynamic systems approach [4], we may conceptualize the role of disregarding the error as follows. Assuming that the behavior of the infant is governed by a dynamic system component for decision making, and another one for execution of movement, the performance-dependent reward signal would be one of the control parameters of the decision making component. In the stable case where behaviors have been learned well (for instance to reach for near objects), negative rewards during exploratory actions would lead to further stabilization of the already learned attractors. If, however, the negative reward is ignored, i.e. the control parameter is changed, existing attractors might be destabilized. This in turn would make it easier for the system to switch to

other attractors, giving their corresponding movements more chance to be practiced in a new context where they would normally not be chosen. Over time, this practice might lead to new stable attractors even under consideration of the error signal once the effect of high dopamine state wears off.

As robots are expected to be active participants in humans daily life, they need to be able to constantly learn and improve their abilities autonomously. The conceptual model and its simplified implementation in the simulation study of this paper offer one possible mechanism for such adaptive behavior acquisition.

VII. Future work

As the preeliminary result with the robot simulator seems to confirm the viability of our approach, the next step in our research is to implement the conceptual model in more detail and evaluate its ability to account for the behavioral data in [1], [2]. Furthermore, we will perform a series of experiments with a real M3-neony humanoid robot, and study the dependence of the results on parameter settings in the simplified version of our model presented in Sec. IV.

VIII. Conclusion

The core idea behind the model was that the level of sense of control determines how much the negative outcome of the action is taken into account for decision making. Omission of the errors was suggested to enable selection of different behaviors in a context when they normally would not be selected providing more learning opportunities for fine-tuning these behaviors.

Acknowledgment

This research was supported by: Ministerio de Ciencia y Innovacion (FPU grant AP2007-02565); Grants-in-Aid for Scientific Research (Research Project Number: 22220002). The authors would like to thank Fabio Dalla Libera for providing the source code of the M3-neony robot simulator.

References

[1] B. J. Grzyb, A. P. del Pobil, and L. B. Smith, "Reaching for the unreachable: (mis)perception of body effectivity in older infants," manuscript in preparation.

[2] ——, "Reaching for the unreachable: the cause or the consequence of learning to walk," manuscript in preparation.

[3] B. J. Grzyb, J. Boedecker, M. Asada, A. P. del Pobil, and L. B. Smith, "Trying anyways: how ignoring the errors may help in learning new skills," in *First Joint IEEE International Conference on Development and Learning and on Epigenetic Robotics*, 2011.

[4] E. Thelen and L. Smith, *A Dynamic Systems Approach to the Development of Cognition and Action*. MIT Press, 1994.

[5] J. D. Cohen, T. S. Braver, and J. W. Brown, "Computational perspectives on dopamine function in prefrontal cortex," *Current Opinion in Neurobiology*, vol. 12(2), pp. 223–229, 2002.

[6] G. A. Barr, S. Moriceau, K. Shionoya, K. Muzny, P. Gao, S. Wang, and R. M. Sullivan, "Transitions in infant learning are modulated by dopamine in the amygdala," *Nature Neuroscience*, vol. 12(11), pp. 1367–1369, 2009.

[7] B. Abler, H. Walter, and S. Erk, "Neural correlates of frustration," *Neuroreport*, vol. 16(7), pp. 669–72, 2005.

[8] W. Schultz, "Multiple reward signals in the brain," *Nature Reviews Neuroscience*, vol. 1, pp. 199–206, 2000.

[9] W. Schultz, P. Dayan, and P. Montague, "A neural substrate of prediction and reward," *Science*, vol. 275 (5306), pp. 1593–1599, 1997.

[10] K. Doya, "Metalearning and neuromodulation," *Neural Networks*, vol. 15(4-6), pp. 495–506, 2002.

[11] R. Sutton, "Learning to predict by the methods of temporal differences," *Machine Learning*, vol. 3(1), pp. 9–44, 1988.

[12] R. S. Sutton and A. G. Barto, *Reinforcement Learning: An Introduction*. MIT Press, Cambridge, MA, 1998.

[13] K. C. Berridge, "The debate over dopamine's role in reward: the case for incentive salience," *Psychopharmacology*, vol. 191, pp. 391–431, 2007.

[14] T. V. Maia and M. J. Frank, "From reinforcement learning models to psychiatric and neurological disorders," *Nature Neuroscience*, vol. 14, pp. 154–162, 2011.

[15] M. R. Delgado, R. H. Frank, and E. A. Phelps, "Perceptions of moral character modulate the neural systems of reward during the trust game," *Nature Neuroscience*, vol. 8, pp. 1611–1618, 2005.

[16] E. Leibenluft, B. A. Rich, D. T. Vinton, E. E. Nelson, S. J. Fromm, L. H. Berghorst, P. Joshi, A. Robb, R. J. Schachar, D. P. Dickstein, E. B. McClure, and D. S. Pine, "Neural circuitry engaged during unsuccessful motor inhibition in pediatric bipolar disorder," *Am J Psychiatry*, vol. 164, pp. 52–60, 2007.

[17] P. Najt, J. Perez, M. Sanches, M. Peluso, D. Glahn, and J. Soares, "Impulsivity and bipolar disorder," *European Neuropsychopharmacology*, vol. 17, pp. 313–320, 2007.

[18] S. Johnson, "Mania and dysregulation in goal pursuit: a review," *Clin Psychol Rev*, vol. 25, pp. 241–262, 2005.

[19] "American psychiatric association (2000). diagnostic and statistical manual of mental disorders. washington, dc."

[20] B. Abler, I. Greenhouse, D. Ongur, H. Walter, and S. Heckers, "Abnormal reward system activation in mania," *Neuropsychopharmacology*, vol. 33, pp. 2217–2227, 2008.

[21] W. Schultz, "The reward signal of midbrain dopamine neurons," *News Physiol. Sci.*, vol. 14, pp. 249–255, 1999.

[22] L. Leotti, S. Iyengar, and K. Ochsner, "Born to choose: the origins and value of the need for control," *Trends Cogn Sci.*, vol. 14(10), pp. 457–463, 2010.

[23] D. J. Shapiro, C. Schwartz, and J. Astin, "Controlling ourselves, controlling our world. psychology's role in understanding positive and negative consequences of seeking and gaining control." *Am Psychol.*, vol. 51(12), pp. 1213–30, 1996.

[24] M. J. Frank, "Hold your horses: A dynamic computational role for subthalamic nucleus in decision making," *Neural Networks*, vol. 19, pp. 1120–1136, 2006.

[25] M. H. Johnson, *Developmental psychology: an advanced textbook*. Psychology Press, 1999, ch. Developmental Neuroscience, pp. 199–230.

[26] H. T. Ghashghaeia, C. C. Hilgetag, and H. Barbasa, "Sequence of information processing for emotions based on the anatomic dialogue between prefrontal cortex and amygdala," *Neuroimage*, vol. 34(3), pp. 905–923, 2007.

[27] M. Milad and G. Quirk, "Neurons in medial prefrontal cortex signal memory for fear extinction," *Nature*, vol. 420(6911), pp. 70–74, 2002.

[28] T. Johnstone, C. M. van Reekum, H. L. Urry, N. H. Kalin, and R. J. Davidson, "Failure to regulate: Counterproductive recruitment of top-down prefrontal-subcortical circuitry in major depression," *The Journal of Neuroscience*, vol. 27(33), pp. 8877–8884, 2007.

[29] P. O'Donnell, *Handbook of basal ganglia structure and function*. Academic Press, 2010, ch. Gating of limbic input to the ventral striatum, pp. 367–37.

[30] R. A. Depue and P. F. Collins, "Neurobiology of the structure of personality: Dopamine, facilitation of incentive motivation, and extraversion," *Behavioral and Brain Sciences*, vol. 22, pp. 491–569, 1999.

[31] M. E. Jackson, A. S. Frost, and B. Moghaddam, "Stimulation of prefrontal cortex at physiologically relevant frequencies inhibits dopamine release in the nucleus accumbens," *Journal of Neurochemistry*, vol. 78, pp. 920–923, 2001.

[32] D. Purves, G. J. Augustine, D. Fitzpatrick, W. C. Hall, A.-S. LaMantia, J. O. McNamara, and L. E. White, Eds., *Neuroscience, Fourth Edition*. Sinauer Associates, Inc., 2007.

[33] R. A. Depue and W. G. Iacono, "Neurobehavioral aspects of affective disorders," *Annu Rev Psychol.*, vol. 40, pp. 457–92, 1989.

[34] A. D. Arco and F. Mora, "Neurotransmitters and prefrontal cortexlimbic system interactions: implications for plasticity and psychiatric disorders," *Journal of Neural Transmission*, vol. 116, pp. 941–952, 2009.