Reward-Punishment Processing and Learning

Hackjin Kim, School of Psychology, Korea University, Seoul, South Korea

Introduction

It has long been debated whether approach and avoidance behaviors are controlled by largely segregated and functionally dissociable neural systems or they are served by a common neural mechanism (Boureauand Dayan, 2011; Cools et al., 2011; Palminteri and Pessiglione, 2017). In fact, much anatomical as well as neuropharmacological evidence seems to support the dual systems view. For example, one network including the ventromedial prefrontal cortex (vmPFC) and the nucleus accumbens (NAC)/the ventral striatum (VS) is preferentially involved in approach behavior, whereas another network including the dorsomedial prefrontal cortex (dmPFC), the dorsal striatum (DS), the insula, and the amygdala is involved in avoidance behavior (Palminteri and Pessiglione, 2017). In addition, this dual systems view seems to be also supported by functional dissociation between dopaminergic (DA) and serotonergic (5-HT) neuromodulatory systems, which have been shown to be involved in approach and avoidance behaviors, respectively (Daw et al., 2002). Alternatively, other theoretical and empirical studies support the unified system view, which suggests that approach and avoidance behaviors are not clearly distinguished from each other, possibly sharing common neural circuitries (Kim et al., 2006; Palminteri et al., 2015; Solomon and Corbit, 1974). In this article, I will first review some key neural substrates of approach and avoidance behaviors, highlighting the inconsistencies between findings that support the dual vs. unified systems views. Later I will suggest an alternative model of approach and avoidance learning based on hierarchical allostatic regulation, whereby conflicts in competing internal bodily needs are regulated by incorporating external sensory information. This model can provide a useful theoretical framework to reconcile the inconsistencies, to integrate the current findings, and to raise concrete and testable hypotheses.

Neural Substrates of Approach and Avoidance Behaviors

Ventral and Dorsal Cortico-Striatal Loop

The ventromedial prefrontal cortex (vmPFC) together with the ventral striatum comprise the ventral cortico-striatal loop (Haber and Knutson, 2010), which has been strongly implicated in approach and reward-seeking behaviors (Bartra et al., 2013; Chib et al., 2009; Delgado et al., 2000; Kable and Glimcher, 2007; Kim et al., 2007; Kim et al., 2011; O'Doherty, 2011). However, the ventral cortico-striatal loop has also been associated with active avoidance behavior (Kim et al., 2006; Levita et al., 2012), encoding successful avoidance particularly under contexts with a negative expected value (Palminteri et al., 2015). Unlike the ventral cortico-striatal loop, its dorsal counterpart including the dmPFC and the dorsal striatum is involved in avoiding punishment (Delgado et al., 2009; Palminteri et al., 2012; Pessiglione et al., 2006; Seymour et al., 2007; Shenhav and Buckner, 2014). The dmPFC has often been shown to respond to negative outcomes such as pain (Rainville et al., 1997), monetary loss (Liu et al., 2011), as well as social rejection (Eisenberger et al., 2003). However, the dorsal cortico-striatal loop has also been implicated in reward learning particularly when effort is required. For example, selective lesioning of dmPFC leads to increased preference for low effort options (Schweimer et al., 2005; Walton et al., 2002), whereas its activity increases positively with expected and actual effort (Scholl et al., 2015; Skvortsova et al., 2014).

There's a growing argument that the vmPFC computes valuation primarily based on internal bodily signals, whereas the dmPFC contributes to valuation based on external sensory information from the environment (Bouretand Richmond, 2010; Howard et al., 2015; Nakao et al., 2012). This argument can be supported by the evidence that the vmPFC is the main target of the dopaminergic projection from the midbrain, which is modulated by either food intake (de Araujo et al., 2012) or direct vagus nerve stimulation (Han et al., 2018). Furthermore, the vmPFC activity is modulated by the experimentally-induced dopamine level (Jocham et al., 2011), visceral signals like hunger and satiety (Howard et al., 2015; Roy et al., 2012), and outcome

devaluation (de Wit et al., 2009; Valentin et al., 2007). Unlike the vmPFC, the dmPFC has been implicated in various psychological functions such as detecting and resolving conflicts among competitive responses (Shenhav et al., 2016), and searching for a new value beyond the current familiar state (Domenech et al., 2020; Kolling et al., 2016). Consistent with these findings, functionally competitive roles between the vmPFC and the dmPFC have been reported. For example, the vmPFC is associated with learning habit-like behaviors, which are formed progressively through overtraining, whereas the dmPFC is responsible for voluntary and goal-directed initial responses (Killcross and Coutureau, 2003). In addition, vmPFC activity increases when agents are more likely to repeat the same decision, whereas dmPFC activity increases when subjective value of the ongoing plan drops and a search for a new plan becomes required (Domenech et al., 2020). Finally, larger and smaller differences in value between two alternative options in a choice task were associated with greater activity in the vmPFC and the dmPFC, respectively (Hackel et al., 2017; Piva et al., 2019). In summary, all the evidence suggests that the vmPFC alone can compute the value of a decision in a familiar or well-practiced task, whereas the dmPFC needs to be engaged to integrate additional information from the environment when there's a conflict to be resolved.

Insula

The anterior insula together with the mid-cingulate cortex has been strongly implicated in negative emotions such as pain (Wager et al., 2013), disgust (Uddin et al., 2017), and financial loss (Kim et al., 2011; Kuhnen and Knutson, 2005). Not surprisingly, it has also been reported to be involved in avoidance behavior. For example, insula activity was shown to correlate with aversive prediction error signals in the avoidance condition (Kim et al., 2006; Pessiglione et al., 2006; Samanez-Larkin et al., 2008; Seymour et al., 2004), and damage to the insula selectively impairs punishment-based avoidance learning, while leaving intact reward-based approach learning (Palminteri et al., 2012). Notably, another line of research has also suggested that the insula along with the mid-cingulate cortex comprise a salience network which is specialized to detect salient changes in the environment and is potentially responsible for behavioral switching (Dajani and Uddin, 2015).

Amygdala

The amygdala has been strongly implicated in aversive learning and avoidance behaviors (De Martino et al., 2010; Mason et al., 2006; Mobbs et al., 2009; Schwartz et al., 2003). However, there is also evidence that the amygdala is involved in both approach and avoidance behaviors, possibly serving a broader role as a behavioral relevance detector (Paton et al., 2006; Sander et al., 2003; Schlund and Cataldo, 2010). In addition, both empirical and theoretical works suggest that the amygdala is related to associative ambiguity or predictive uncertainty (Hsu et al., 2005; Whalen, 2007).

Dopamine Versus Serotonin Functions

Dopamine (DA) is well known to be one of the major neurotransmitters related to reward learning and approach behaviors (Cohen et al., 2012; Matsumoto and Hikosaka, 2009; Pessiglione et al., 2006; Schultz, 1998). Unlike DA, serotonin (5-HT) has been mainly associated with avoidance behaviors (Crockett et al., 2009; Evers et al., 2005; Geurts et al., 2013; Macoveanu, 2014), and its primary functions appear to be roughly opposite to those of DA (Boureau and Dayan, 2011; Cools et al., 2011; Daw et al., 2002). However, there is also evidence that 5-HT can be involved in approach behavior, especially when animals have been trained to wait for a reward (Kranz et al., 2010; Li et al., 2016). For example, optogenetic activation of 5-HT neurons promoted waiting for an expected reward in a delayed reward task, suggesting their role in the reward anticipatory behavior (Fonseca et al., 2015; Miyazaki et al., 2011; Miyazaki et al., 2014; Schweighofer et al., 2008; Zhou et al., 2015). Such a reward anticipatory function of 5-HT may be related to its more general function of inhibiting perseverative responding and promoting cognitive flexibility (Matias et al., 2017). Supporting this idea, it was shown that 5-HT could contribute to cognitive flexibility by directly suppressing both impulsive and perseverative responses through behavioral inhibition (Boureau and Dayan, 2011; Cools et al., 2011).

5-HT can decrease DA-induced behavior-reinforcing effect, potentially balancing the effects of DA (Fischer and Ullsperger, 2017), and the modulatory effects of DA on 5-HT neurons are often weaker than those of 5-HT on DA neurons (Cools et al., 2011). Such a modulatory effect of 5-HT on DA can also be supported at least partly by the differences between them in anatomical connectivity. First, the habenula is known to increase 5-HT release and inhibit DA neurons, and has been strongly implicated in aversive learning (Hikosaka, 2010). It is noteworthy, however, that habenula lesion impairs active avoidance learning, but not the defensive behavior induced by classical fear conditioning (Amo et al., 2010). In addition, the habenula plays a key role in switching behavioral strategies by integrating new evidence with prior experience (Palumbo et al., 2020). Second, DA projections from the ventral tegmental area are rather confined to the ventral cortico-striatal loops (Murty et al., 2014), whereas serotonergic projections are more extensive to include more dorsal cortico-striatal loops as well as the saliency network including the insula and the anterior mid-cingulate cortex (Bar et al., 2016). These differences in network connectivity between DA and 5-HT can be supported by both animal and human studies. For example, aggression is associated with increased DA level in the ventral striatum and decreased 5-HT level in the mPFC in rats (van Erp and Miczek, 2000), and

the degree to which dorsal striatal activity correlated with expected future rewards was enhanced by an increase in 5-HT level, whereas the degree to which ventral striatal activity correlated with expected immediate rewards was enhanced by a decrease in 5-HT level in humans (Tanaka et al., 2007).

In summary, although approach and avoidance behaviors appear to be served by distinctive neural circuitries and neuromodulatory systems, many mixed findings exist. In fact, these neural features associated with approach and avoidance seem more likely to reflect the degree to which external sensory information is required to address internal bodily needs as an agent needs to update pre-existing behavioral strategy. For example, the ventral cortico-striatal loop and DA system may be specialized for reinforcing previously successful behavioral strategy, whereas the dorsal loop, the saliency network, and 5-HT system may be involved in suppressing previously failed response and searching for a new strategy. This can be related to the long-held idea of stability-plasticity (Grossberg, 1980) or exploitation-exploration dilemma (Daw et al., 2006). To elaborate this view, I will introduce a hierarchical allostatic regulation model of valuation below.

Hierarchical Allostatic Regulation Model of Approach and Avoidance Learning

Homeostasis can be a major drive underlying a wide range of behaviors in vertebrates (Cannon, 1939; Craig, 2003). In this sense, both approach and avoidance can be viewed as different subtypes of behavioral strategies to address the imbalance of bodily homeostasis. For example, an animal may approach food to address reduced glucose level in the blood and may avoid a pain-inducing stimulus to maintain homeostatic state in bodily tissue. Therefore, from this point of view, it can be argued that approach and avoidance behaviors serve the common goals, that are, preventing and maintaining bodily homeostasis.

If both approach and avoidance are simply different means of achieving the same goal, then how is it that we often witness seemingly segregated neural features distinctively related to each of them? The answer may lie in the fundamental asymmetry in the strategy and the amount of external sensory information required for achieving the goals. For example, a reward can be a sign that you may stay in the current state or repeat what you just did, whereas a punishment can be a sign that you need to find other alternative actions than what you just did (Hikosaka, 2010). Such an asymmetry could be in parallel with the well-known decision heuristic of *win-stay lose-shift*, which has been proven evolutionarily advantageous and more powerful than other complex rules (Nowak and Sigmund, 1993). According to this strategy, an agent repeats the previous action if the outcome was satisfying and seeks alternative actions otherwise, dealing with failure vs. success in a completely different manner. Therefore, at least parts, if not all, of the neural differences between approach and avoidance listed above may reflect this fundamental asymmetry between them in terms of strategy for determining complexity and type of information to be processed.

Recent theoretical works propose hierarchical allostatic regulation models of brain function for computing values of decision, which can explain how our brain creates, maintains, and updates values of decision in various situations (Kim, 2020; Smith et al., 2017). According to these models, the brain consists of multiple functionally dissociable layers or levels that are hierarchically organized such that the lower levels are more sensitive to internal signals from bodily organs and the higher levels are more sensitive to external signals from the environment. When conflicts occur at the lower levels, the upper levels would be engaged to resolve them by increasing sensitivity to incoming external sensory signals from the environment. The upper level can take over decision control temporarily to search for a new and more sophisticated value of decision. The upper level then sends prediction signals to update the pre-existing value at the lower level and continues this cycle of search until it finds a new value that can effectively resolve the conflict at the lower level (Fig. 1).

This model can provide an alternative and perhaps more comprehensive account of why the ventral and dorsal cortico-striatal loops are often engaged in approach and avoidance behavior, respectively. For example, unlike approach behaviors, avoidance behaviors would require greater amounts of external information because an agent needs to search for other options except for the one just selected and invalidated, which would then require an engagement of the dorsal loop that is more involved in processing additional external information. From this perspective, one can speculate that experiencing negative outcome may trigger neural processes of re-allocating attention to the environment to search for a new, potentially better, alternative, whereas experiencing positive outcome may elicit a simpler strategy of maintaining previously chosen behavior that have led to a successful consequence. According to this model, the insula and the amygdala, which have been implicated in avoidance behavior, also contribute to detecting conflicts and activating external valuation to resolve the conflicts based on additional external sensory information. Furthermore, 5-HT may additionally recruit the external valuation system, which would then interact with the internal system to resolve the conflicts, whereas DA influences are rather confined to the internal valuation system. Such a hierarchical nature of the difference between approach and avoidance could also provide a more comprehensive account for why insula damage, which affects the external valuation, impairs avoidance learning but not approach learning (Palminteri et al., 2012), whereas vmPFC damage, which affects the internal valuation, can abolish both approach and avoidance learning (Camille et al., 2011; Wheeler and Fellows, 2008) because the engagement of the external valuation requires conflicts in the internal valuation.
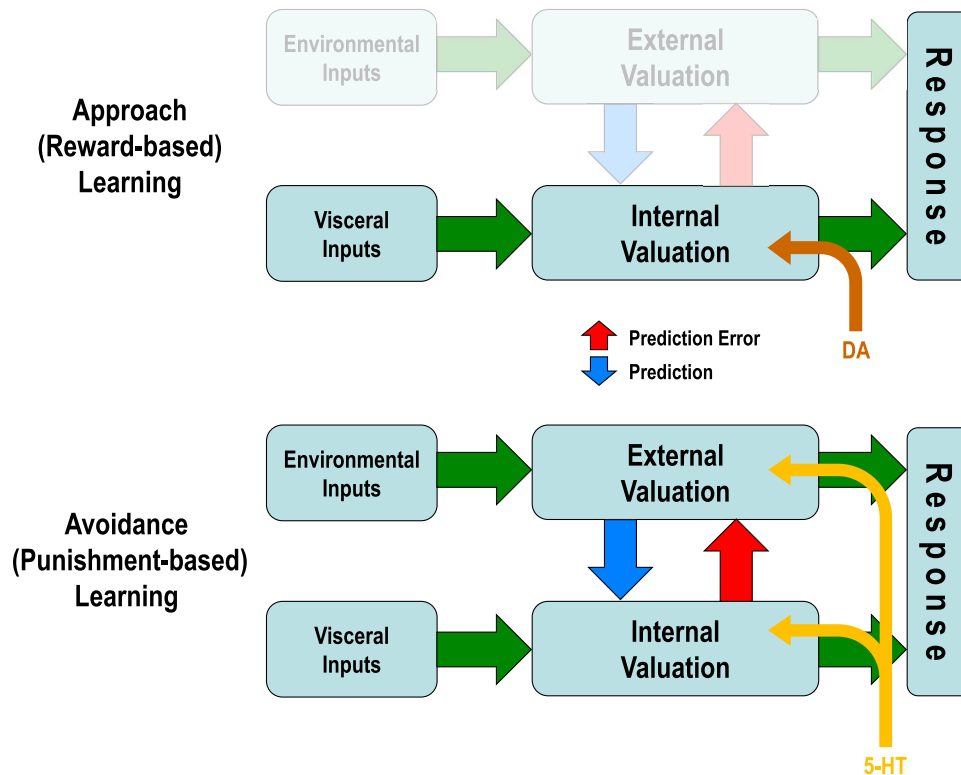
Fig. 1 A Hierarchical Allostatic Regulation Model for Approach and Avoidance Learning. The model assumes that the brain comprises functionally dissociable hierarchically organized valuation systems, which are differentially involved in computing values of decision along the ventral-to-dorsal spatial gradient of increasing external sensory inputs and decreasing internal visceral inputs. The *internal valuation* system, which includes the ventral cortico-striatal loop, generates interoceptive prediction signals and elicits a familiar intuitive response to prevent foreseen bodily imbalance based on past experiences. When two or more mutually incompatible values are simultaneously activated at the internal valuation system, a conflict (prediction error) can occur. Such a conflict can be detected by the saliency system including the insula, which can switch the currently engaged decision mode from the internal valuation system to the *external valuation* system including the dorsal cortico-striatal loop. The external valuation system would then increase the sensitivity to incoming sensory signals from the external environment to resolve the conflict, sending prediction signals to update the pre-existing value encoded at the lower level. This process continues until it finds a new value that resolves the conflict, and the newly updated value will be strengthened and internalized through repetition so that it is activated more quickly and easily in similar situations later without causing a conflict. This model also assumes that dopamine neurons in the ventral tegmental area selectively project to the internal valuation system, whereas serotonergic raphe nuclei project more broadly to both internal and the external valuation systems.

Conclusions and Future Directions

In this article, it was suggested that both approach and avoidance behaviors serve the common goals of survival ultimately aimed at maintaining bodily homeostasis, and that the asymmetry between approach and avoidance behaviors may arise from the degree to which the two types of behavior differ in terms of utilizing external information from the environment. An alternative model of approach and avoidance behaviors, based on the hierarchical allostatic regulation model, was proposed, which predicts that different levels of the cortico-striatal loops would be engaged as a function of the amount of external information that needs to be processed, regardless of whether they are approach or avoidance. This model is expected to present many testable hypotheses and questions that could spur research in new directions to improve our understanding of the neural networks involved in approach and avoidance behaviors.

References

Amo, R., Aizawa, H., Takahoko, M., Kobayashi, M., Takahashi, R., Aoki, T., Okamoto, H., 2010. Identification of the zebrafish ventral habenula as a homolog of the mammalian lateral habenula. J. Neurosci. 30 (4), 1566–1574. https://doi.org/10.1523/JNEUROSCI.3690-09.2010.
Bar, K.J., de la Cruz, F., Schumann, A., Koehler, S., Sauer, H., Critchley, H., Wagner, G., 2016. Functional connectivity and network analysis of midbrain and brainstem nuclei. NeuroImage 134, 53–63. https://doi.org/10.1016/j.neuroimage.2016.03.071.
Bartra, O., McGuire, J.T., Kable, J.W., 2013. The valuation system: a coordinate-based meta-analysis of BOLD fMRI experiments examining neural correlates of subjective value. NeuroImage 76, 412–427. https://doi.org/10.1016/j.neuroimage.2013.02.063.

Boureau, Y.L., Dayan, P., 2011. Opponency revisited: competition and cooperation between dopamine and serotonin. Neuropsychopharmacology 36 (1), 74–97. https://doi.org/10.1038/npp.2010.151.

Bouret, S., Richmond, B.J., 2010. Ventromedial and orbital prefrontal neurons differentially encode internally and externally driven motivational values in monkeys. J. Neurosci. 30 (25), 8591–8601. https://doi.org/10.1523/JNEUROSCI.0049-10.2010.

Camille, N., Griffiths, C.A., Vo, K., Fellows, L.K., Kable, J.W., 2011. Ventromedial frontal lobe damage disrupts value maximization in humans. J. Neurosci. 31 (20), 7527–7532. https://doi.org/10.1523/JNEUROSCI.6527-10.2011.

Cannon, W.B., 1939. The Wisdom of the Body. WW Norton and Company, New York.

Chib, V.S., Rangel, A., Shimojo, S., O'Doherty, J.P., 2009. Evidence for a common representation of decision values for dissimilar goods in human ventromedial prefrontal cortex. J. Neurosci. 29 (39), 12315–12320. https://doi.org/10.1523/JNEUROSCI.2575-09.2009.

Cohen, J.Y., Haesler, S., Vong, L., Lowell, B.B., Uchida, N., 2012. Neuron-type-specific signals for reward and punishment in the ventral tegmental area. Nature 482 (7383), 85–88. https://doi.org/10.1038/nature10754.

Cools, R., Nakamura, K., Daw, N.D., 2011. Serotonin and dopamine: unifying affective, activational, and decision functions. Neuropsychopharmacology 36 (1), 98–113. https://doi.org/10.1038/npp.2010.121.

Craig, A.D., 2003. A new view of pain as a homeostatic emotion. Trends Neurosci. 26 (6), 303–307. https://doi.org/10.1016/s0166-2236(03)00123-1.

Crockett, M.J., Clark, L., Robbins, T.W., 2009. Reconciling the role of serotonin in behavioral inhibition and aversion: acute tryptophan depletion abolishes punishment-induced inhibition in humans. J. Neurosci. 29 (38), 11993–11999. https://doi.org/10.1523/JNEUROSCI.2513-09.2009.

Dajani, D.R., Uddin, L.Q., 2015. Demystifying cognitive flexibility: implications for clinical and developmental neuroscience. Trends Neurosci. 38 (9), 571–578. https://doi.org/10.1016/j.tins.2015.07.003.

Daw, N.D., Kakade, S., Dayan, P., 2002. Opponent interactions between serotonin and dopamine. Neural Netw. 15 (4–6), 603–616. https://doi.org/10.1016/s0893-6080(02)00052-7.

Daw, N.D., O'Doherty, J.P., Dayan, P., Seymour, B., Dolan, R.J., 2006. Cortical substrates for exploratory decisions in humans. Nature 441 (7095), 876–879. https://doi.org/10.1038/nature04766.

de Araujo, I.E., Ferreira, J.G., Tellez, L.A., Ren, X., Yeckel, C.W., 2012. The gut-brain dopamine axis: a regulatory system for caloric intake. Physiol. Behav. 106 (3), 394–399. https://doi.org/10.1016/j.physbeh.2012.02.026.

De Martino, B., Camerer, C.F., Adolphs, R., 2010. Amygdala damage eliminates monetary loss aversion. Proc. Natl. Acad. Sci. U. S. A. 107 (8), 3788–3792. https://doi.org/10.1073/pnas.0910230107.

de Wit, S., Corlett, P.R., Aitken, M.R., Dickinson, A., Fletcher, P.C., 2009. Differential engagement of the ventromedial prefrontal cortex by goal-directed and habitual behavior toward food pictures in humans. J. Neurosci. 29 (36), 11330–11338. https://doi.org/10.1523/JNEUROSCI.1639-09.2009.

Delgado, M.R., Nystrom, L.E., Fissell, C., Noll, D.C., Fiez, J.A., 2000. Tracking the hemodynamic responses to reward and punishment in the striatum. J. Neurophysiol. 84 (6), 3072–3077. https://doi.org/10.1152/jn.2000.84.6.3072.

Delgado, M.R., Jou, R.L., Ledoux, J.E., Phelps, E.A., 2009. Avoiding negative outcomes: tracking the mechanisms of avoidance learning in humans during fear conditioning. Front. Behav. Neurosci. 3, 33. https://doi.org/10.3389/neuro.08.033.2009.

Domenech, P., Rheims, S., Koechlin, E., 2020. Neural mechanisms resolving exploitation-exploration dilemmas in the medial prefrontal cortex. Science 369 (6507). https://doi.org/10.1126/science.abb0184.

Eisenberger, N.I., Lieberman, M.D., Williams, K.D., 2003. Does rejection hurt? An FMRI study of social exclusion. Science 302 (5643), 290–292. https://doi.org/10.1126/science.1089134.

Evers, E.A., Cools, R., Clark, L., van der Veen, F.M., Jolles, J., Sahakian, B.J., Robbins, T.W., 2005. Serotonergic modulation of prefrontal cortex during negative feedback in probabilistic reversal learning. Neuropsychopharmacology 30 (6), 1138–1147. https://doi.org/10.1038/sj.npp.1300663.

Fischer, A.G., Ullsperger, M., 2017. An update on the role of serotonin and its interplay with dopamine for reward. Front. Hum. Neurosci. 11, 484. https://doi.org/10.3389/fnhum.2017.00484.

Fonseca, M.S., Murakami, M., Mainen, Z.F., 2015. Activation of dorsal raphe serotonergic neurons promotes waiting but is not reinforcing. Curr. Biol. 25 (3), 306–315. https://doi.org/10.1016/j.cub.2014.12.002.

Geurts, D.E., Huys, Q.J., den Ouden, H.E., Cools, R., 2013. Serotonin and aversive Pavlovian control of instrumental behavior in humans. J. Neurosci. 33 (48), 18932–18939. https://doi.org/10.1523/JNEUROSCI.2749-13.2013.

Grossberg, S., 1980. How does a brain build a cognitive code? Psychol. Rev. 87 (1), 1–51. https://doi.org/10.1007/978-94-009-7758-7_1.

Haber, S.N., Knutson, B., 2010. The reward circuit: linking primate anatomy and human imaging. Neuropsychopharmacology 35 (1), 4–26. https://doi.org/10.1038/npp.2009.129.

Hackel, L.M., Zaki, J., Van Bavel, J.J., 2017. Social identity shapes social valuation: evidence from prosocial behavior and vicarious reward. Soc. Cogn. Affect Neurosci. 12 (8), 1219–1228. https://doi.org/10.1093/scan/nsx045.

Han, W., Tellez, L.A., Perkins, M.H., Perez, I.O., Qu, T., Ferreira, J., et al., 2018. A neural circuit for gut-induced reward. Cell 175 (3), 887–888. https://doi.org/10.1016/j.cell.2018.10.018.

Hikosaka, O., 2010. The habenula: from stress evasion to value-based decision-making. Nat. Rev. Neurosci. 11 (7), 503–513. https://doi.org/10.1038/nrn2866.

Howard, J.D., Gottfried, J.A., Tobler, P.N., Kahnt, T., 2015. Identity-specific coding of future rewards in the human orbitofrontal cortex. Proc. Natl. Acad. Sci. U. S. A. 112 (16), 5195–5200. https://doi.org/10.1073/pnas.1503550112.

Hsu, M., Bhatt, M., Adolphs, R., Tranel, D., Camerer, C.F., 2005. Neural systems responding to degrees of uncertainty in human decision-making. Science 310 (5754), 1680–1683. https://doi.org/10.1126/science.1115327.

Jocham, G., Klein, T.A., Ullsperger, M., 2011. Dopamine-mediated reinforcement learning signals in the striatum and ventromedial prefrontal cortex underlie value-based choices. J. Neurosci. 31 (5), 1606–1613. https://doi.org/10.1523/JNEUROSCI.3904-10.2011.

Kable, J.W., Glimcher, P.W., 2007. The neural correlates of subjective value during intertemporal choice. Nat. Neurosci. 10 (12), 1625–1633.

Killcross, S., Coutureau, E., 2003. Coordination of actions and habits in the medial prefrontal cortex of rats. Cerebr. Cortex 13 (4), 400–408. Retrieved from. http://www.ncbi.nlm.nih.gov/pubmed/12631569.

Kim, H., Shimojo, S., O'Doherty, J.P., 2006. Is avoiding an aversive outcome rewarding? Neural substrates of avoidance learning in the human brain. PLoS Biol. 4 (8), e233. https://doi.org/10.1371/journal.pbio.0040233.

Kim, H., Adolphs, R., O'Doherty, J.P., Shimojo, S., 2007. Temporal isolation of neural processes underlying face preference decisions. Proc. Natl. Acad. Sci. U. S. A. 104 (46), 18253–18258.

Kim, H., Shimojo, S., O'Doherty, J.P., 2011. Overlapping responses for the expectation of juice and money rewards in human ventromedial prefrontal cortex. Cerebr. Cortex 21 (4), 769–776. https://doi.org/10.1093/cercor/bhq145.

Kim, H., 2020. Stability or plasticity? - a hierarchical allostatic regulation model of medial prefrontal cortex function for social valuation. Front. Neurosci. 14 https://doi.org/10.3389/fnins.2020.00281. ARTN 281.

Kolling, N., Wittmann, M.K., Behrens, T.E.J., Boorman, E.D., Mars, R.B., Rushworth, M.F.S., 2016. Value, search, persistence and model updating in anterior cingulate cortex. Nat. Neurosci. 19 (10), 1280–1285. https://doi.org/10.1038/nn.4382.

Kranz, G.S., Kasper, S., Lanzenberger, R., 2010. Reward and the serotonergic system. Neuroscience 166 (4), 1023–1035. https://doi.org/10.1016/j.neuroscience.2010.01.036.

Kuhnen, C.M., Knutson, B., 2005. The neural basis of financial risk taking. Neuron 47 (5), 763–770. https://doi.org/10.1016/j.neuron.2005.08.008.

Levita, L., Hoskin, R., Champi, S., 2012. Avoidance of harm and anxiety: a role for the nucleus accumbens. NeuroImage 62 (1), 189–198. https://doi.org/10.1016/j.neuroimage.2012.04.059.

Li, Y., Zhong, W., Wang, D., Feng, Q., Liu, Z., Zhou, J., et al., 2016. Serotonin neurons in the dorsal raphe nucleus encode reward signals. Nat. Commun. 7, 10503. https://doi.org/10.1038/ncomms10503.

Liu, X., Hairston, J., Schrier, M., Fan, J., 2011. Common and distinct networks underlying reward valence and processing stages: a meta-analysis of functional neuroimaging studies. Neurosci. Biobehav. Rev. 35 (5), 1219–1236. https://doi.org/10.1016/j.neubiorev.2010.12.012.

Macoveanu, J., 2014. Serotonergic modulation of reward and punishment: evidence from pharmacological fMRI studies. Brain Res. 1556, 19–27. https://doi.org/10.1016/j.brainres.2014.02.003.

Mason, W.A., Capitanio, J.P., Machado, C.J., Mendoza, S.P., Amaral, D.G., 2006. Amygdalectomy and responsiveness to novelty in rhesus monkeys (Macaca mulatta): generality and individual consistency of effects. Emotion 6 (1), 73–81. https://doi.org/10.1037/1528-3542.6.1.73.

Matias, S., Lottem, E., Dugue, G.P., Mainen, Z.F., 2017. Activity patterns of serotonin neurons underlying cognitive flexibility. Elife 6. https://doi.org/10.7554/eLife.20552.

Matsumoto, M., Hikosaka, O., 2009. Two types of dopamine neuron distinctly convey positive and negative motivational signals. Nature 459 (7248), 837–841. https://doi.org/10.1038/nature08028.

Miyazaki, K., Miyazaki, K.W., Doya, K., 2011. Activation of dorsal raphe serotonin neurons underlies waiting for delayed rewards. J. Neurosci. 31 (2), 469–479. https://doi.org/10.1523/JNEUROSCI.3714-10.2011.

Miyazaki, K.W., Miyazaki, K., Tanaka, K.F., Yamanaka, A., Takahashi, A., Tabuchi, S., Doya, K., 2014. Optogenetic activation of dorsal raphe serotonin neurons enhances patience for future rewards. Curr. Biol. 24 (17), 2033–2040. https://doi.org/10.1016/j.cub.2014.07.041.

Mobbs, D., Marchant, J.L., Hassabis, D., Seymour, B., Tan, G., Gray, M., et al., 2009. From threat to fear: the neural organization of defensive fear systems in humans. J. Neurosci. 29 (39), 12236–12243. https://doi.org/10.1523/JNEUROSCI.2378-09.2009.

Murty, V.P., Shermohammed, M., Smith, D.V., Carter, R.M., Huettel, S.A., Adcock, R.A., 2014. Resting state networks distinguish human ventral tegmental area from substantia nigra. NeuroImage 100, 580–589. https://doi.org/10.1016/j.neuroimage.2014.06.047.

Nakao, T., Ohira, H., Northoff, G., 2012. Distinction between externally vs. Internally guided decision-making: operational differences, meta-analytical comparisons and their theoretical implications. Front. Neurosci. 6, 31. https://doi.org/10.3389/fnins.2012.00031.

Nowak, M.A., Sigmund, K., 1993. A strategy of win-stay, lose-shift that outperforms tit-for-tat in the prisoner's dilemma game. Nature 364 (6432), 56–58. https://doi.org/10.1038/364056a0.

O'Doherty, J.P., 2011. Contributions of the ventromedial prefrontal cortex to goal-directed action selection. Ann. N. Y. Acad. Sci. 1239, 118–129. https://doi.org/10.1111/j.1749-6632.2011.06290.x.

Palminteri, S., Pessiglione, M., 2017. Opponent brain systems for reward and punishment learning: causal evidence from drug and lesion studies in humans. In: Dreher, J.C., Tremblay, L. (Eds.), Decision Neuroscience. Academic Press, pp. 291–303.

Palminteri, S., Justo, D., Jauffret, C., Pavlicek, B., Dauta, A., Delmaire, C., et al., 2012. Critical roles for anterior insula and dorsal striatum in punishment-based avoidance learning. Neuron 76 (5), 998–1009. https://doi.org/10.1016/j.neuron.2012.10.017.

Palminteri, S., Khamassi, M., Joffily, M., Coricelli, G., 2015. Contextual modulation of value signals in reward and punishment learning. Nat. Commun. 6, 8096. https://doi.org/10.1038/ncomms9096.

Palumbo, F., Serneels, B., Pelgrims, R., Yaksi, E., 2020. The zebrafish dorsolateral habenula is required for updating learned behaviors. Cell Rep. 32 (8), 108054. https://doi.org/10.1016/j.celrep.2020.108054.

Paton, J.J., Belova, M.A., Morrison, S.E., Salzman, C.D., 2006. The primate amygdala represents the positive and negative value of visual stimuli during learning. Nature 439 (7078), 865–870. https://doi.org/10.1038/nature04490.

Pessiglione, M., Seymour, B., Flandin, G., Dolan, R.J., Frith, C.D., 2006. Dopamine-dependent prediction errors underpin reward-seeking behaviour in humans. Nature 442 (7106), 1042–1045. https://doi.org/10.1038/nature05051.

Piva, M., Velnoskey, K., Jia, R., Nair, A., Levy, I., Chang, S.W., 2019. The dorsomedial prefrontal cortex computes task-invariant relative subjective value for self and other. Elife 8. https://doi.org/10.7554/eLife.44939.

Rainville, P., Duncan, G.H., Price, D.D., Carrier, B., Bushnell, M.C., 1997. Pain affect encoded in human anterior cingulate but not somatosensory cortex. Science 277 (5328), 968–971. https://doi.org/10.1126/science.277.5328.968.

Roy, M., Shohamy, D., Wager, T.D., 2012. Ventromedial prefrontal-subcortical systems and the generation of affective meaning. Trends Cogn. Sci. 16 (3), 147–156. https://doi.org/10.1016/j.tics.2012.01.005.

Samanez-Larkin, G.R., Hollon, N.G., Carstensen, L.L., Knutson, B., 2008. Individual differences in insular sensitivity during loss anticipation predict avoidance learning. Psychol. Sci. 19 (4), 320–323. https://doi.org/10.1111/j.1467-9280.2008.02087.x.

Sander, D., Grafman, J., Zalla, T., 2003. The human amygdala: an evolved system for relevance detection. Rev. Neurosci. 14 (4), 303–316. https://doi.org/10.1515/revneuro.2003.14.4.303.

Schlund, M.W., Cataldo, M.F., 2010. Amygdala involvement in human avoidance, escape and approach behavior. NeuroImage 53 (2), 769–776. https://doi.org/10.1016/j.neuroimage.2010.06.058.

Scholl, J., Kolling, N., Nelissen, N., Wittmann, M.K., Harmer, C.J., Rushworth, M.F., 2015. The good, the bad, and the irrelevant: neural mechanisms of learning real and hypothetical rewards and effort. J. Neurosci. 35 (32), 11233–11251. https://doi.org/10.1523/JNEUROSCI.0396-15.2015.

Schultz, W., 1998. Predictive reward signal of dopamine neurons. J. Neurophysiol. 80 (1), 1–27. https://doi.org/10.1152/jn.1998.80.1.1.

Schwartz, C.E., Wright, C.I., Shin, L.M., Kagan, J., Rauch, S.L., 2003. Inhibited and uninhibited infants "grown up": adult amygdalar response to novelty. Science 300 (5627), 1952–1953. https://doi.org/10.1126/science.1083703.

Schweighofer, N., Bertin, M., Shishida, K., Okamoto, Y., Tanaka, S.C., Yamawaki, S., Doya, K., 2008. Low-serotonin levels increase delayed reward discounting in humans. J. Neurosci. 28 (17), 4528–4532. https://doi.org/10.1523/JNEUROSCI.4982-07.2008.

Schweimer, J., Saft, S., Hauber, W., 2005. Involvement of catecholamine neurotransmission in the rat anterior cingulate in effort-related decision making. Behav. Neurosci. 119 (6), 1687–1692. https://doi.org/10.1037/0735-7044.119.6.1687.

Seymour, B., O'Doherty, J.P., Dayan, P., Koltzenburg, M., Jones, A.K., Dolan, R.J., et al., 2004. Temporal difference models describe higher-order learning in humans. Nature 429 (6992), 664–667. https://doi.org/10.1038/nature02581.

Seymour, B., Daw, N., Dayan, P., Singer, T., Dolan, R., 2007. Differential encoding of losses and gains in the human striatum. J. Neurosci. 27 (18), 4826–4831. https://doi.org/10.1523/JNEUROSCI.0400-07.2007.

Shenhav, A., Buckner, R.L., 2014. Neural correlates of dueling affective reactions to win-win choices. Proc. Natl. Acad. Sci. U. S. A. 111 (30), 10978–10983. https://doi.org/10.1073/pnas.1405725111.

Shenhav, A., Cohen, J.D., Botvinick, M.M., 2016. Dorsal anterior cingulate cortex and the value of control. Nat. Neurosci. 19 (10), 1286–1291. Retrieved from <Go to ISI>://WOS:000384549900008.

Skvortsova, V., Palminteri, S., Pessiglione, M., 2014. Learning to minimize efforts versus maximizing rewards: computational principles and neural correlates. J. Neurosci. 34 (47), 15621–15630. https://doi.org/10.1523/JNEUROSCI.1350-14.2014.

Smith, R., Thayer, J.F., Khalsa, S.S., Lane, R.D., 2017. The hierarchical basis of neurovisceral integration. Neurosci. Biobehav. Rev. 75, 274–296. https://doi.org/10.1016/j.neubiorev.2017.02.003.

Solomon, R.L., Corbit, J.D., 1974. An opponent-process theory of motivation. I. Temporal dynamics of affect. Psychol. Rev. 81 (2), 119–145. https://doi.org/10.1037/h0036128.

Tanaka, S.C., Schweighofer, N., Asahi, S., Shishida, K., Okamoto, Y., Yamawaki, S., Doya, K., 2007. Serotonin differentially regulates short- and long-term prediction of rewards in the ventral and dorsal striatum. PLoS One 2 (12), e1333. https://doi.org/10.1371/journal.pone.0001333.

Uddin, L.Q., Nomi, J.S., Hebert-Seropian, B., Ghaziri, J., Boucher, O., 2017. Structure and function of the human insula. J. Clin. Neurophysiol. 34 (4), 300–306. https://doi.org/10.1097/WNP.0000000000000377.

Valentin, V.V., Dickinson, A., O'Doherty, J.P., 2007. Determining the neural substrates of goal-directed learning in the human brain. J. Neurosci. 27 (15), 4019–4026. https://doi.org/10.1523/JNEUROSCI.0564-07.2007.

van Erp, A.M., Miczek, K.A., 2000. Aggressive behavior, increased accumbal dopamine, and decreased cortical serotonin in rats. J. Neurosci. 20 (24), 9320–9325. Retrieved from. https://www.ncbi.nlm.nih.gov/pubmed/11125011.

Wager, T.D., Atlas, L.Y., Lindquist, M.A., Roy, M., Woo, C.W., Kross, E., 2013. An fMRI-based neurologic signature of physical pain. N. Engl. J. Med. 368 (15), 1388–1397. https://doi.org/10.1056/NEJMoa1204471.

Walton, M.E., Bannerman, D.M., Rushworth, M.F., 2002. The role of rat medial frontal cortex in effort-based decision making. J. Neurosci. 22 (24), 10996–11003. Retrieved from. https://www.ncbi.nlm.nih.gov/pubmed/12486195.

Whalen, P.J., 2007. The uncertainty of it all. Trends Cogn. Sci. 11 (12), 499–500. https://doi.org/10.1016/j.tics.2007.08.016.

Wheeler, E.Z., Fellows, L.K., 2008. The human ventromedial frontal lobe is critical for learning from negative feedback. Brain 131 (Pt 5), 1323–1331. https://doi.org/10.1093/brain/awn041.

Zhou, J., Jia, C., Feng, Q., Bao, J., Luo, M., 2015. Prospective coding of dorsal raphe reward signals by the orbitofrontal cortex. J. Neurosci. 35 (6), 2717–2730. https://doi.org/10.1523/JNEUROSCI.4017-14.2015.