Dopamine neurons coding prediction errors in reward space, but not in aversive space: a matter of location?

Joachim Morrens
Graduate Student Cognitive and Decision Sciences, Division of Psychology and Language Sciences, Faculty of Brain Sciences, University College London, London, United Kingdom

Submitted 21 October 2013; accepted in final form 24 February 2014

A PROMINENT FRAMEWORK CALLED reinforcement learning theory describes how individuals can learn by experience, which future decisions they should make so as to maximize rewards and minimize punishments (Daw and Tobler 2014). By the coding of reward prediction errors (RPEs), dopamine neurons of the midbrain play a central role in the neural implementation of this learning theory.

A RPE can be described as the difference between the observed state of the world (i.e., what you get) and the previously expected state of the world (i.e., what you thought you would get). For instance, when an individual consumes a good that turns out to be better than anticipated, this gives rise to a positive RPE. The positive RPE is accompanied by release of dopamine which in turn presumably plays a role in the biophysical implementation of learning in brain areas receiving the dopamine projections (Daw and Tobler 2014; Schultz 2013). Similarly, individuals who have learned to associate a reward-predicting cue with a reward, display increased dopamine neuron activity at the usual time of reward in case the reward is unexpectedly omitted, a so-called negative RPE.

Both positive and negative RPEs involve the engagement of dopamine neurons in a context of reward, hence the name RPEs. A fundamental question is, however, whether dopamine neurons of the midbrain are also involved in the processing of aversive stimuli, that is, whether dopamine neurons operate in the same way in aversive space as they have been shown to do in reward space.

This question was recently addressed by Fiorillo (2013), and it is this study that makes up the focus of this Neuro Forum article. Specifically, Fiorillo attempted to distinguish two value dimension hypotheses with respect to activity of dopamine neurons. According to the single dimension hypothesis, appetitiveness and aversiveness are two opposite sides of a single continuous value dimension, or value space. Consequently, this theory predicts dopamine neurons to signal prediction errors (PEs) related to aversive stimuli in the same manner as they do for appetitive stimuli. Alternatively, the two dimensions hypothesis claims that appetitiveness and aversiveness make up two separate, discrete value dimensions. Ergo, this hypothesis suspects one class of neurons (e.g., dopamine neurons) to operate in reward space, while another operates in aversive space.

To distinguish between both hypotheses, Fiorillo (2013) analyzed single neuron recording data of 195 dopamine neurons from two rhesus macaque monkeys. Neurons were recorded during simple Pavlovian tasks involving both appetitive and aversive stimuli: orally delivered juice, orally delivered saline and bitter solutions, and airpuffs delivered to the nose. A prior choice task was used to equalize the absolute subjective value of appetitive and aversive stimuli. This was done to ensure that the absolute subjective value of the aversive stimuli was high enough, such that they would certainly have an influence on dopamine neuron activity (Fiorillo et al. 2013).

In a first experiment where a conditioned stimulus predicted subsequent reward with a chance of 0.5, reward delivery elicited increased dopamine neuron activity, while reward omission suppressed dopamine neuron activity. This finding confirms many earlier reports (see Daw and Tobler 2014), and it is in line with both value dimension hypotheses. However, the single dimension hypothesis projects that aversive or neutral stimuli which are not as bad as (“better than”) expected should also lead to a positive RPE and should thus evoke increased dopamine neuron activity. Intriguingly, this is exactly what was not observed by Fiorillo: dopamine firing still “dipped” at the time of aversive or neutral stimulus delivery, even when those stimuli were not as bad as anticipated (see Fig. 1 in Fiorillo 2013). This finding provides a first indication that dopamine neurons behave markedly different in aversive space compared with reward space.

In a second experiment, Fiorillo traced the effect of cues announcing appetitive and aversive stimuli. In line with the traditional RPE hypothesis, appetitive stimuli only elicited increased firing of dopamine neurons when they came by surprise. Prediction of aversive stimuli, in contrast, had only a minor effect on suppression of dopamine neuron activity: a more or less equal decrease in dopamine neuron activity was
Supplementary Fig. 9
mine neurons when announced airpuffs were omitted (see their and Hikosaka (2009), who did in fact report excitation of dopamine neurons. However, it is in contrast with an earlier finding from Matsumoto and Hikosaka (2009), who did in fact report excitation of dopamine neurons when announced airpuffs were omitted (see their Supplementary Fig. 9C). Likewise, these authors showed that dopamine neurons are less suppressed when airpuffs are fully anticipated (see their Fig. 3C). A potential explanation for these contrasting findings is one of location. This is because of the gradient that dopamine neurons in the midbrain exhibit: the more ventral in the midbrain one goes, that is, the more toward the ventral tegmental area (VTA) and the ventral tier of the substantia nigra pars compacta (SNc), the more dopamine neuron activity reflects the coding of RPEs (Matsumoto and Hikosaka 2009). Conversely, the more dorsal in the midbrain one goes, the more dopamine neurons tend to get activated by both appetitive and aversive events, consistent with the coding of motivational salience-related signals (Matsumoto and Hikosaka 2009; see also Horvitz 2000). Given the existence of this gradient, it is important to note that, in contrast to Matsumoto and Hikosaka’s data, Fiorillo’s recordings mainly stemmed from neurons in the dorsal SNc, with relatively few VTA neurons included. Notwithstanding this potential issue of recording location, Fiorillo’s observed dopamine responses also do not match the predictions of dorsal motivational salience coding (see Matsumoto and Hikosaka 2009), which would require a positive response to the aversive stimuli in experiment 1 and 2. In any case, whether the current observations would also hold for dopamine neurons that are specifically located in the VTA, seems an interesting question.

In Fiorillo’s third and final experiment, monkeys had to learn the meaning of several predictive cues: those that announced 1) orally delivered juice, 2) orally delivered juice in combination with an orally delivered saline/bitter solution, and 3) orally delivered juice in combination with an airpuff to the nose.

It was observed that a predictive cue, which announced juice alone, caused stronger increased activity in dopamine neurons than a predictive cue announcing juice in combination with saline or bitter solution. This result seems in line with both value dimension hypotheses; as the saline/bitter solution can be expected to directly diminish the value of the juice, it makes sense that the expected reward is not as large as it would have been in the juice alone condition. In this case, dopamine neurons are thus capable of effectively integrating expected benefits and costs in reward space.

However, a cue predicting juice plus airpuff caused, compared with the juice alone condition, suppression of dopamine neuron firing in monkey F, but not in monkey O (see Fig. 1 and compare Fig. 3B and Fig. 3C in Fiorillo 2013). Fiorillo nonetheless described this result as “only a small suppression” of dopamine neuron activity in monkey F and further noted that the suppressing effect of saline was significantly greater than that of airpuff. An extra experiment was conducted in which a neutral sound of 90 dB replaced the airpuff, and also announcement of the concurrent neutral sound was ineffective in suppressing dopamine neuron firing. All the same, this additional experiment was only performed in monkey O, the one that already did not show any sign of suppression during the initial airpuff experiment.

Fiorillo (2013) concluded that airpuff does not cause strong suppression of dopamine neuron firing, an interpretation which is, according to the author, fully in line with the two dimensions hypothesis: a predicted aversive stimulus alters dopamine firing only when it can be expected to directly affect the value of the appetitive stimulus. According to this hypothesis, prediction of saline and bitter, orally delivered together with juice, significantly diminishes the value of the announced juice reward. In contrast, the simultaneous delivery of airpuff to the nose has little effect on the anticipated reward value of the juice itself and is therefore less effective in devaluing the predicted juice reward. Yet, Fiorillo subtly ignored the fact that one monkey’s dopamine neurons did in fact seem to integrate the negative value of the airpuff with the positive value of the juice. So, the following question emerges: Why is it that one monkey’s neurons exhibited this behavior, while the other monkey’s neurons did not?

One potential explanation for this finding is, again, the dopamine subpopulation gradient mentioned earlier (Matsumoto and Hikosaka 2009), which was also found in a more recent study by the same research group (Matsumoto and Takada 2013). This is because slightly different neuronal subpopulations were recorded from both monkeys with the majority of ventrally recorded neurons belonging to monkey F (see Fig. 1 in Fiorillo et al. 2013). This raises the question whether monkey F’s sample contained more “real” RPE neurons. In any case, the differences in recording location seem, at first sight, not extreme enough to fully account for this contrasting observation, and the heterogeneity of dopamine midbrain neurons is believed to be gradual rather than discrete (e.g., Fiorillo et al. 2013; Matsumoto and Takada 2013). As already hinted above, it would have nonetheless been informative if Fiorillo had given us data on responses of dopamine neurons subdivided by their anatomical location.

A second potential explanation is one of different subjective experience. By now, it is well established that dopamine neurons encode the subjective, rather than the objective, experience of stimuli (Schultz 2013). One can therefore speculate that monkey F perceived the announced airpuff as if it would directly “spoil” his anticipated juice reward, while monkey O could better separate the values of the two announced stimuli. Therefore, it would be interesting for future work to explore similar potential idiosyncrasy of integration of costs and benefits in dopamine neuron responses.

Fiorillo (2013) finishes his discussion by stating that the present results point to the existence of four types of value representations, denoted as reward-ON, reward-OFF, aversive-ON, and aversive-OFF. While the ON-neurons are activated by evidence for reward or for aversiveness, the OFF-neurons are activated by evidence against reward or against aversiveness. Fiorillo claims that dopamine neurons embody reward-ON neurons and infers that the current results imply the existence of...
of three other major modulatory neurotransmitter systems which represent the three other value representations. While such framework is certainly a possibility, I believe the present results are also compatible with the existence of only two systems: one system computing PEs in reward space (the dopamine system) and another one computing PEs in aversive space (currently unknown). An often recurring remark, which may be the reason why Fiorillo considers dopamine neurons not to be involved with the reward-OFF representation, is that a positive RPE involves a phasic tripling of firing rates, while a negative RPE has very limited dynamic range due to the low baseline activity (0.1–7 Hz) of dopamine neurons (see Daw and Tobler 2014). However, Hart et al. (2014) recently demonstrated that actual dopamine concentrations, measured with fast-scan cyclic voltammetry in the nucleus accumbens of rats, are consistent with a bidirectional RPE signal involving symmetrical coding of positive and negative RPEs. In contrast to Fiorillo’s claim, their results thus support the idea that dopamine encodes the full range of PEs necessary for reinforcement learning in reward space.

Fiorillo’s (2013) present contribution adds to a body of evidence which insinuates that appetitive and aversive stimuli are processed by, at least partly, separate neurotransmitter systems (e.g., Boureau and Dayan 2011). Nevertheless, it should be kept in mind that Fiorillo’s dominant recording location might have contained little true RPE neurons, and previous work did in fact show some evidence of aversive PE coding in such neurons (Matsumoto and Hikosaka 2009). Additionally, some studies have even found dopamine neurons to be involved in fear-related learning (e.g., Tan et al. 2012). Stating that dopamine is definitely not involved in learning in aversive space seems, therefore, still a premature conclusion.

ACKNOWLEDGMENTS
I thank Sebastian Haesler for fruitful discussions.

DISCLOSURES
No conflicts of interest, financial or otherwise, are declared by the author(s).

AUTHOR CONTRIBUTIONS
J.M. drafted manuscript; J.M. edited and revised manuscript; J.M. approved final version of manuscript.

J Neurophysiol • doi:10.1152/jn.00751.2013 • www.jn.org
REFERENCES


