Received: 20 August 2019 | Accepted: 10 September 2019

DOI: 10.1111/ejn.14578

FEATURED PAPER COMMENTARY

EJN European Journal of Neuroscience FENS WILEY

Uncertainty versus prediction error in Pavlovian fear conditioning: Commentary on Walker et al. (2019)

Natalie C. Tronson

Department of Psychology, University of Michigan, Ann Arbor, MI, USA

Correspondence
Natalie C. Tronson, Department of Psychology, University of Michigan, Ann Arbor, MI 48109, USA.
Email: ntronson@umich.edu

Prediction error, or a mismatch between what is expected and what occurs, is a fundamental feedback mechanism across modalities including sensory, attention, motivation, and associative learning. Identifying what has changed in our environment allows us to appropriately direct attention toward new information. In learning tasks, "surprise"—or the unexpected occurrence or omission of a biologically relevant event—is thought to act as a teaching signal that drives new learning. Neural correlates of prediction error provide additional support for a central role in learning. Most famously, striatal dopaminergic neurons fire during an unexpected reward, but not to a well-predicted reinforcer (Schultz, Dayan, & Montague, 1997). Yet, these "prediction error" signals may alternatively result from motivation or attentional processes (Nasser, Calu, Schoenbaum, & Sharpe, 2017), suggesting that prediction is not the only information learned Pavlovian conditioning.

In their paper in this issue of European Journal of Neuroscience, Walker, Wright, Jhou, and McDannald (2019) examine the role of ventrolateral periaqueductal gray (vlPAG) during fear conditioning. Specifically, they observed increased firing in vlPAG to a footshock unconditioned stimulus (US) after partial reinforcement training, compared with a continuous reinforcement condition in which the US is well predicted. This suggests that increased firing in response to footshock during uncertainty is a result of prediction error, and not due to the sensory properties or salience of the US. Second, they show that inhibition of vlPAG results in a decrease in fear-responding in subsequent trials, suggesting that prediction error is necessary for maintaining fear-responding during uncertainty. Overall, this is a neat set of findings that fits more-or-less within established narratives for prediction error and learning (Fernández, Boccia, & Pedreira, 2016). These findings, however, also highlight uncertainties about prediction and prediction error in Pavlovian conditioning. In particular, what kinds of prediction are generated after partial reinforcement, and what other processes drive conditioned responding when the outcome is uncertain?

The assumption that animals learn to predict the occurrence of a biologically salient event (the US) stems in large part from the role of prediction error as a teaching signal in quantitative learning theories (McNally, Johansen, & Blair, 2011). Under conditions of continuous reinforcement where the US reliably follows the CS, organisms learn to predict the US based from the presence of the CS, resulting in conditioned responses triggered by the CS—increasing the ability to retrieve food or avoid danger. But after partial reinforcement, the outcome is uncertain: the CS is sometimes, but not always, followed by the US. What, then, do animals predict?

One possibility is that animals learn a probabilistic prediction of US occurrence. If the CS-US contingency is 0.375, then animals learn to predict that the probability of receiving the US on any given trial is $p = 0.375$. How, then, does this trigger prediction error? If the US occurrence either occurs ($p = 1$) or is omitted ($p = 0$), then does every trial generate a prediction error? If the animal is already expecting a US on some trials, is the US ever surprising? Or is there an accumulation of expectancy across trials, like waiting for a bus, with each non-reinforced trial increasing expectancy of shock after the next CS presentation (Glimcher, 2011)? An alternative to this probabilistic schema, animals could learn that the CS may or may not be followed by the US—and then generate a stochastic guess about US occurrence. Without a prediction about US occurrence, neither occurrence nor omission of the US would generate an error signal.

There are several reasons to doubt a purely predictive account of partial reinforcement. First, after partial reinforcement, performance does not reflect a strict probabilistic prediction. Fear-related behaviors as measured either by conditioned suppression (Walker *et al.*, 2019) or by freezing (Huh *et al.*, 2009) are stronger than expected based simply

on reinforcement probability. This suggests that multiple sources of information, in addition to prediction, drive conditioned fear responses. Second, there is strong evidence for a failure to generate prediction errors after omission of the US. Conditioned responses after partial reinforcement are resistant to extinction after fear conditioning, and animals fail to generate negative prediction error signals that precede context fear extinction (Huh *et al.*, 2009). This suggests that animals become insensitive to the CS-US contingency after partial reinforcement and that factors other than prediction drive conditioned responding.

One largely overlooked contribution to fear conditioning and defensive responses is motivational processes. Recent work has demonstrated the importance of learned affective and motivational information in Pavlovian conditioning. In appetitive Pavlovian conditioning, for example, organisms attribute motivational salience to the CS, resulting in intense wanting based on the value of the US representation and the current state of the animal (Berridge, 2018). Here, motivational or affective information is learned in parallel to predictive information, and factors including individual differences (Cogan, Shapses, Robinson, & Tronson, 2019) and uncertainty (Anselme, 2010) determine whether predictive or motivational components of Pavlovian associations drive behavior.

Motivational and emotion-related information likely play analogous roles in fear conditioning (Berridge, 2018). Further, there is evidence that motivational processes are critical to process uncertain biologically or psychologically significant events (Anselme, 2010) and drive increased anxiety under partial reinforcement (Grillon, Baas, Cornwell, & Johnson, 2006). Previous work in fear conditioning supports this idea, demonstrating that motivational or affective learning to a fear CS increases after partial reinforcement when prediction is poor (Huh *et al.*, 2009). Interestingly, partial but not continuous reinforcement also increases salience of an aversive US (Hall, Prados, & Sansa, 2005), suggesting that uncertainty plays a critical role in modulating motivational aspects of both the CS, and the US representation.

Motivation or salience-related information may also explain the vlPAG error-like signaling observed in the present study. Because vlPAG also plays multiple roles in complex sensory perception, pain, and antinociception, as well as salience and fear learning (McNally *et al.*, 2011), the observed vlPAG may contribute to any one of these processes. For example, increased US salience might contribute to the increased neuronal response to a US after partial but not continuous reinforcement observed here (Walker *et al.*, 2019). Decreased US salience, rather than disruption of prediction error, may also contribute to reduced future fear-responding after inhibition of vlPAG after partial (but not continuous) reinforcement.

Here, we suggest that Pavlovian fear memories are multidimensional constructs, with elements including motivation, stimulus representation, and value, as well as prediction. The relative strength of each memory component varies as a function of training parameters (e.g., prediction will be stronger after continuous than partial reinforcement), and influenced by factors including individual differences, sex differences, prior experience, and current motivational state. By careful experimental dissociation of these processes, we can begin to unpack the multiple contributions to vulnerability to dysregulation of memory and pathological fear.

ORCID

Natalie C. Tronson https://orcid.org/0000-0001-5676-1579

REFERENCES

Anselme, P. (2010). The uncertainty processing theory of motivation. *Behavioral Brain Research*, *208*, 291–310.

Berridge, K. C. (2018). Evolving concepts of emotion and motivation. *Frontiers in Psychology*, *9*, 1647.

Cogan, E. S., Shapses, M. A., Robinson, T. E., & Tronson, N. C. (2019). Disrupting reconsolidation: memory erasure or blunting of emotional/motivational value? *Neuropsychopharmacology*, *44*, 399–407.

Fernández, R. S., Boccia, M. M., & Pedreira, M. E. (2016). The fate of memory: Reconsolidation and the case of Prediction Error. *Neuroscience and Biobehavioral Reviews*, *68*, 423–441.

Glimcher, P. W. (2011). Understanding dopamine and reinforcement learning: the dopamine reward prediction error hypothesis. *Proceedings of the National Academy of Sciences of the United States of America*, *108*(Suppl 3), 15647–15654.

Grillon, C., Baas, J. M. P. P., Cornwell, B., & Johnson, L. (2006). Context conditioning and behavioral avoidance in a virtual reality environment: Effect of predictability. *Biological Psychiatry*, *60*, 752–759.

Hall, G., Prados, J., & Sansa, J. (2005). Modulation of the effective salience of a stimulus by direct and associative activation of its representation. *Journal of Experimental Psychology: Animal Behavior Processes*, *31*, 267–276.

Huh, K. H., Guzman, Y. F., Tronson, N. C., Guedea, A. L., Gao, C., & Radulovic, J. (2009). Hippocampal Erk mechanisms linking prediction error to fear extinction: Roles of shock expectancy and contextual aversive valence. *Learning & Memory*, *16*, 273–278.

McNally, G. P., Johansen, J. P., & Blair, H. T. (2011). Placing prediction into the fear circuit. *Trends in Neurosciences*, *34*, 283–292.

Nasser, H. M., Calu, D. J., Schoenbaum, G., & Sharpe, M. J. (2017). The dopamine prediction error: Contributions to associative models of reward learning. *Frontiers in Psychology*, *8*, 244.

Schultz, W., Dayan, P., & Montague, P. R. (1997). A neural substrate of prediction and reward. *Science*, *275*, 1593–1598.

Walker, R. A., Wright, K. M., Jhou, T. C., & McDannald, M. A. (2019). The ventrolateral periaqueductal gray updates fear via positive prediction error. *European Journal of Neuroscience*. https://doi.org/10.1111/ejn.14536 [Epub ahead of print].