Trends Cogn Sci. Author manuscript; available in PMC 2021 April 01.

Published in final edited form as:
Trends Cogn Sci. 2020 April ; 24(4): 259–260. doi:10.1016/j.tics.2019.12.005.

Predicting to Perceive and Learning when to Learn

Philip Corlett
Yale University, 34 Park St., CNRU, New Haven, CT, 06519, United States

In their insightful piece, Press et al. scrutinize predictive processing theories of perception [1], which hold that perception involves inferences based on prior beliefs and prediction errors: in some cases, priors explain away incoming data, facilitating the processing of surprising events. In other cases, events consistent with priors are afforded extra processing. How can this be?

The authors suggest that on initial apprehension, features consistent with priors receive preferential processing, however, subsequently, cancellation reigns and prediction errors garner more processing. There are data that are consistent with this account. However, there are problems too. For example, Press et al. demand that inferences consistent with priors mediate 'veridical' perception. This seems to contradict theory and observation with regards to perceptual illusions, wherein the inference to the best explanation is (i) driven by priors and (ii) deviates from the actual input [2]. Furthermore, the proposed temporal order (augmentation of predicted features followed by cancellation), could place surprises too late to be adaptive (it would seem prudent to respond as swiftly as possible to a tiger in my living room), and does not honor the pre-emptive cancellation of self-generated sensory consequences of movements (of the eyes, the head, speech musculature) embodied in corollary discharge theories of motor control and the attribution of agency. Corollary discharges are not preceded by a period of enhancement which is what Press and colleagues would predict.

These issues aside, the authors are right to focus on the inconsistencies in theory and data around predictive processing and perception. Their allusion to learning reminds me of similar paradoxes in formal animal learning theory. In what remains, I adumbrate an account of learning and belief updating that might compliment the authors' and ultimately reconcile some of the incongruent observations.

Learning has long been implicated in perception – Pavlov remarked; 'what the genius Helmholtz was referring to in unconscious inference is the mechanism of the conditioned reflex' [3]. Rescorla and Wagner (R&W) evoked prediction error to account for conditioning effects like Kamin blocking, wherein mere contiguity is insufficient and surprising changes in contingency drive learning [4]. In humans, blocking occurs for causal and social inferences, but also in low level perceptual phenomena like contingent color after effects [5]. However, associability, the proclivity for cues and outcomes to garner learning, is fixed in the R&W model. Various phenomena in rodents (and humans) suggest that prediction errors

Correspondence: philip.corlett@yale.edu.

change associability, however, echoing Press et al.'s observation, the theories disagree on the direction of change – in some circumstances we learn most from predictive cues [6], in others we learn most when uncertainty is greatest [7].

There are data consistent with both theories. Statistical learning models [8] might reconcile these data. Such models simultaneously track the relevance of cues for predicting outcomes (their reliability [7]), and the their uncertainty, or failure to correctly predict outcomes, which adjusts subsequent predictions [6] (sometimes these qualities are referred to as expected and unexpected uncertainty, respectively). An apple tree is a reliable predictor of apples, relative to other trees, but nonetheless it may still be an uncertain one (given the seasons or weather). In rodents, lesions of medial prefrontal cortex and parietal cortex doubly dissociate reliability-based from uncertainty-based attention [9]. I propose that reliability based predictive learning – about actions – is facilitated by modeling the impact of oneself as an agent, and, via predictive cancellation, discerning whether or not one was the cause of some salient event in the world [10]. Uncertainty-based inference on the other hand is key for updating associations with new learning, which, barring some catastrophic events, should not usually be required for the consequences of one's own action. Thus, reliability mechanisms should be more strongly involved in actions and their impact on perception, whereas uncertainty mechanisms may be involved more broadly, when a model of the self is less critical (e.g. when learning about external environmental events or agents), or when the self-model fails and needs updating with new learning [11].

The extrapolation from conditioning preparations to perceptual paradigms demands careful thought (for example, what serves as cue versus outcome?), however, initial results are encouraging. The dopaminergic prediction error signal – a stalwart in the neurobiology of reinforcement learning, similarly underwrites learning of associations between sensory events [12], as one would predict if these learning mechanisms were not concerned with value per se, but rather the causal texture of the world, of which our bodies and actions are key parts. When this learning errs, the symptoms of serious mental illness arise. They may provide empirical contexts for exploring prior beliefs and the inputs they explain away. Initial work seems consistent with the idea that unreliable predictions about oneself may bias inferences toward external cues, relatively unrelated to the self [11].

Acknowledgements

I thank Pantelis Leptourgos for discussions. Any errors that remain are mine. I was supported by the Yale University Department of Psychiatry, the Connecticut Mental Health Center (CMHC) and Connecticut State Department of Mental Health and Addiction Services (DMHAS), and by an IMHRO / Janssen Rising Star Translational Research Award, an Interacting Minds Center (Aarhus) Pilot Project Award, NIMH R01MH12887, and R21MH120799

References

1. Press C et al. (2019) The Perceptual Prediction Paradox. Trends Cogn Sci.
2. Geisler WS and Kersten D (2002) Illusions, perception and Bayes. Nat Neurosci 5 (6), 508–10. [PubMed: 12037517]
3. Pavlov IP (1928) Natural science and the brain In Lectures on conditioned reflexes: Twenty-five years of objective study of the higher nervous activity (behaviour) of animals, Liverwright Publishing Corporation.

4. Rescorla RA, Wagner AR (1972) A theory of Pavlovian conditioning: Variations in the effectiveness of reinforcement and non-reinforcement In Classical conditioning II: Current research and theory (Black AH, Prokasy WF ed), Appleton-Century-Crofts.
5. Westbrook RF and Harrison W (1984) Associative blocking of the McCollough effect. The Quarterly journal of experimental psychology. A, Human experimental psychology 36 (2), 309–18. [PubMed: 6533689]
6. Pearce JM and Hall G (1980) A model for Pavlovian learning: variations in the effectiveness of conditioned but not of unconditioned stimuli. Psychol Rev 87 (6), 532–52. [PubMed: 7443916]
7. Mackintosh NJ (1975) A theory of attention: Variations in the associability of stimuli with reinforcement. Psychological Review 82 (276–298).
8. Dayan P et al. (2000) Learning and selective attention. Nat Neurosci 3 Suppl, 1218–23. [PubMed: 11127841]
9. Holland PC and Schiffino FL (2016) Mini-review: Prediction errors, attention and associative learning. Neurobiol Learn Mem 131, 207–15. [PubMed: 26948122]
10. Redgrave P and Gurney K (2006) The short-latency dopamine signal: a role in discovering novel actions? Nat Rev Neurosci 7 (12), 967–75. [PubMed: 17115078]
11. Corlett PR et al. (2019) Hallucinations and Strong Priors. Trends Cogn Sci 23 (2), 114–127. [PubMed: 30583945]
12. Sharpe MJ et al. (2017) Dopamine transients are sufficient and necessary for acquisition of model-based associations. Nat Neurosci 20 (5), 735–742. [PubMed: 28368385]