Gold et al. (1) report a reinforcement-learning experiment where reward prediction errors (RPEs) were elicited by the consonance or dissonance of musical stimuli. They link these RPEs to activation in the nucleus accumbens (NAc) and to behavioral indices of learning. They conclude that music can function as a reward, that musical expectations are linked to pleasure, and that musical rewards motivate learning. We applaud the multifaceted methodological approach, which combines neuroimaging, behavioral data and computational modeling. However, we believe that some of the conclusions put forward are not warranted by the evidence presented.

First, the paper conflates sensory prediction (predicting future events) with reward prediction (predicting rewards for future events) (2, 3). Fundamentally, the study addresses reward prediction, yet the abstract defines RPEs as signaling “the difference between expected and perceived musical events” (i.e., sensory prediction), and the introduction motivates the experiment by citing prior research on sensory prediction. While sensory prediction enjoys a long tradition of research linking it to musical aesthetics (4, 5), there is currently little evidence that reward prediction plays such an important role. Conflating the 2 types of prediction is problematic, because it implies that investigating reward prediction sheds light on the aesthetic role of sensory prediction, which is not the case.

Furthermore, while the paper purports to investigate musical RPEs, reward predictions in the experiment came from visual rather than musical cues, so it is not clear that the results generalize to typical music listening, where predictions (whether for reward or sensory events) primarily arise from the musical structure itself. This limits the ecological validity of the paradigm, which is better interpreted as a generic visual reinforcement learning task, where the rewarding stimulus simply happens to be musical.

Lastly, the individual-differences analysis is undermined by methodological issues. The authors use stepwise regression to compensate for the multicollinearity in their features, yet stepwise regression (already heavily criticized in the statistical literature) is particularly vulnerable to multicollinearity (6–8). Moreover, the authors use robust regression to protect from the outliers in Figures 4A and 4B, yet these are high-leverage outliers, to which robust regression (specifically Least Absolute Deviation) is particularly vulnerable (9). If the outliers are removed, these associations lose statistical significance, implying that the association between right NAc activity and learning was considerably less clear than originally argued. Reproducible code for these reanalyses is available at https://doi.org/10.24433/CO.4493464.v1.
To conclude, while this study shows that music can function as a rewarding and reinforcing stimulus, it tells us little about the pleasures of ecological music listening because it addresses reward prediction in the absence of sensory prediction, using visual cues that are typically absent from music listening, and with statistical analyses that are driven by high-leverage outliers. However, we do think that musical reward prediction deserves further exploration, and we look forward to the authors’ future contributions in this area.