Subjective, more than objective, expectation and surprise explain perceptual decisions during learning

Jessye Clarke [1], Kirsten Rittershofer [1,2], Daniel Yon [2], Clare Press [1]

1. Department of Experimental Psychology and Functional Imaging Laboratory, University College London, WC1 H 0AP.
2. School of Psychological Sciences, Birkbeck, University of London, WC1E 7HX

Abstract

Under popular ‘predictive coding’ accounts in cognitive neuroscience, the brain continuously generates predictions about sensory input and integrates them with incoming signals in order to form a percept. There is evidence that these perceptual predictions and corresponding prediction errors can be entirely implicit. However, we are selectively aware of some violations of objective statistical structure - triggering conscious experiences of surprise. In order to investigate the influence of this subjective awareness on learning and perception, we conducted two behavioural studies pairing a probabilistic perceptual discrimination task with trial-by-trial ratings of subjective expectation (Experiment 1) and surprise (Experiment 2). We found that the subjective experience associated with predictions and prediction errors can explain independent variance in behaviour to that explained by the ‘objective’ expectedness or prediction error parameter of a learning model. This suggests that beyond just the presence of a statistically probable or improbable event, subjective awareness of statistical regularities or prediction errors influences downstream stimulus processing and behavioural responses during perception and learning.

Introduction

‘Bayesian’ theories of perception posit that our perceptual system increases the accuracy of perception - on average - by weighting incoming information in line with a previously perceived statistical norm. If the sensory input is misaligned with the expected outcome, this generates a prediction error, which can trigger the updating of an internal predictive model (den Ouden, Kok & de Lange, 2012; de Lange, Heilbron & Kok, 2018). Experimental evidence suggests that the statistical learning enabling such perceptual predictions can occur in the absence of awareness (Turk-Browne et al, 2009), and that corresponding prediction errors may be processed, prompt neural responses, and influence learning without any experience of a mismatch between expectation and reality (Czigler et al, 2007; Rowe, Tsuchiya & Garrido, 2022). However, other violations of statistical regularities are broadcast to awareness, often with subjective experiences of surprise. As most previous work investigating the role of expectations on perception has relied on definitions of expected and unexpected predefined by the probabilistic task structure, it remains unknown whether subjective experiences of objective statistical structure - or correspondingly, experiences of violated expectations or surprise when these statistical relationships are violated - shape how sensory information is processed, perceived and acted upon. In two experiments we asked if subjective expectation (Experiment 1) or surprise (Experiment 2) associated with a cued stimulus could explain independent variance in perceptual decisions, relative to that explained by a Rescorla-Wagner learning model.

Procedure

In both experiments, participants were trained in 75% contingency mappings between two audio tones and two orientated Gabor patches (Fig.1). On each trial, participants were asked to discriminate the orientation of the Gabor patch, and to subsequently rate how much they had expected this orientation (Experiment 1) or how surprising they had found it (Experiment 2).

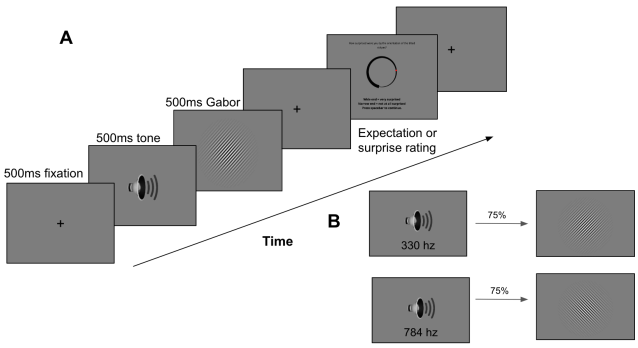

Figure 1. Panel A: Example trial structure. Panel B: Example tone-orientation mapping.

Modelling

To derive ‘objective’ values for stimulus expectedness and prediction error, we used a Rescorla-Wagner (RW) learning model (Rescorla & Wagner, 1972). RW parameters have been shown to explain behaviour (Roesch et al, 2012; Williams et al, 2017) and brain responses (Rodriguez, Aron & Poldrack, 2006; den Ouden et al, 2008; Roesch et al, 2012) during learning. The model maintains separate expectation values (VT) for each tone, initialized at 0 and updated on each trial using $VT(t) = VT(t-1) + \alpha(\gamma - VT(t-1))$, where $VT(t)$ represents the model's expectation for the given tone at trial t, and γ is the presented orientation,

coded as 1 or -1. On each trial, the model computes a prediction error value using $PE = \gamma - VT(t - 1)$.
In Experiment 1, we derived a measure of expectation strength independent of stimulus orientation by multiplying counterclockwise expectations by -1. We adjusted for 25% condition trials by computing $(VT * -1)$. In Experiment 2 we used the absolute prediction error directly as our 'objective' values.

Results

Experiment 1. In Experiment 1 (n=35), participants discriminated the orientation of a cued Gabor and rated how much they had expected the presented orientation on each trial. Responses were quicker ($t(34) = -6.36$, $p < 0.001$) and more accurate ($t(34) = 2.83$, $p = 0.008$) for expected stimuli, indicating successful learning of tone-orientation associations. Linear mixed effects and logistic regression models were fitted predicting reaction time and accuracy respectively from expectedness ratings, RW associative strengths or both. All reaction time models had significant negative fixed effects ($p < 0.001$), indicating that as expectedness increased, reaction times decreased. For accuracy, subjective and objective expectation models had significant positive fixed effects ($p < 0.001$), while the full model showed a significant positive effect of subjective ($p < 0.001$) but not model expectations ($p = 0.082$). BIC scores showed that the best fitting model of both accuracy and RTs was the subjective model (RT BIC = 5316, accuracy BIC = 4117), outperforming the objective model (RT BIC = 5325, accuracy BIC = 4135) and the full model (RT BIC = 5321, accuracy BIC = 4126) for both behavioural measures. This suggests that subjective expectations associated with cued stimuli capture additional and unique behavioural variance to 'objective' RW statistical expectations.

Experiment 2. In Experiment 2 (n=36), we paired the same experimental design with trial-wise ratings of surprise. Participants again responded faster ($t(35) = -5.73$, $p < 0.001$) and more accurately ($t(35) = 2.84$, $p = 0.007$) to expected stimuli. All reaction time models had significant effects ($p < 0.001$), indicating a positive association between both subjective surprise and objective prediction error and response time. For accuracy, subjective and objective models had fixed negative effects ($p < 0.001$), while the full model had a negative effect for surprise ($p < 0.001$) and a nonsignificant negative effect of prediction errors ($p = 0.11$). In terms of model fit, as in Experiment 1 the subjective surprise model (RT BIC = 2764, accuracy BIC = 6164), outperformed the full (RT BIC = 2772, accuracy BIC = 6172) and the objective (RT BIC = 2770, accuracy BIC = 6188) models in both behavioural measures. This indicates that subjective experiences of surprise associated with perceptual prediction errors influence behaviour during learning, and this influence cannot be described solely by the magnitude of 'objective' model generated prediction errors.

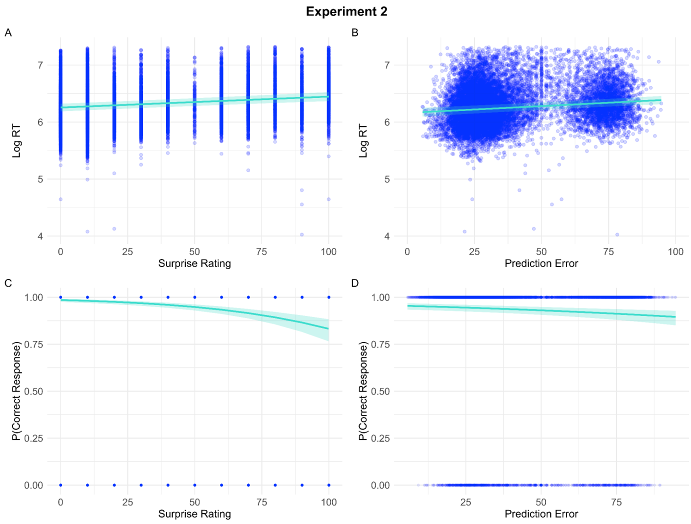

Figure 2. Fixed effects for subjective surprise (left panel) and objective prediction error (right panel) models predicting reaction time and accuracy in Experiment 2. 95% confidence intervals displayed in turquoise for each model.

Discussion

In this work, we investigated the role of subjective expectation and surprise associated with predictions and prediction errors in a probabilistic perceptual discrimination task. We demonstrated that subjective experiences of objective statistical structure - and violations of this structure - explain significant variance in behaviour during learning. We further found that this variance is different, and in some cases greater, to that explained by objective task structure itself, modelled via a RW learning model. This suggests that while objective statistical associations and prediction errors can be processed implicitly and still influence learning, awareness of this structure via subjective expectation or surprise plays an additional role in influencing downstream stimulus processing, perception, and behavioural responses.

This work was supported by a European Research Council (ERC) consolidator grant (101001592) under the European Union’s Horizon 2020 research and innovation programme, both awarded to CP.

References

de Lange, F. P., Heilbron, M., & Kok, P. (2018). How Do Expectations Shape Perception? *Trends in Cognitive Sciences*, *22*(9), 764–779. https://doi.org/10.1016/j.tics.2018.06.002

den Ouden, H. E. M., Friston, K. J., Daw, N. D., McIntosh, A. R., & Stephan, K. E. (2008). A Dual Role for Prediction Error in Associative Learning. *Cerebral Cortex*, *19*(5), 1175–1185. https://doi.org/10.1093/cercor/bhn161

den Ouden, H. E. M., Kok, P., & de Lange, F. P. (2012). How prediction errors shape perception, attention, and motivation. *Frontiers in Psychology*, *3*(548). https://doi.org/10.3389/fpsyg.2012.00548

Rescorla, R. A., & Wagner, A. R. (1972). A theory of Pavlovian conditioning: Variations in the effectiveness of reinforcement and nonreinforcement. In A. H. Black & W. F. Prokasy (Eds.), *Classical Conditioning II: Current Research and Theory* (pp. 64–99). New York: Appleton-Century-Crofts.

Rodriguez, P. F., Aron, A. R., & Poldrack, R. A. (2006). Ventral–striatal/nucleus–accumbens sensitivity to prediction errors during classification learning. *Human Brain Mapping*, *27*(4), 306–313. https://doi.org/10.1002/hbm.20186

Roesch, M. R., Esber, G. R., Li, J., Daw, N. D., & Schoenbaum, G. (2012). Surprise! Neural correlates of Pearce-Hall and Rescorla-Wagner coexist within the brain. *European Journal of Neuroscience*, *35*(7), 1190–1200. https://doi.org/10.1111/j.1460-9568.2011.07986.x

Rowe, E. G., Tsuchiya, N., & Garrido, M. I. (2020). Detecting (Un)seen Change: The Neural Underpinnings of (Un)conscious Prediction Errors. *Frontiers in Systems Neuroscience*, *14*. https://doi.org/10.3389/fnsys.2020.541670

Turk-Browne, N. B., Scholl, B. J., Chun, M. M., & Johnson, M. K. (2009). Neural Evidence of Statistical Learning: Efficient Detection of Visual Regularities Without Awareness. *Journal of Cognitive Neuroscience*, *21*(10), 1934–1945. https://doi.org/10.1162/jocn.2009.21131