Characterizing the mechanisms of instructed reinforcement learning with fMRI pattern analysis

Euan Prentis, Nathan Tardiff, & Sharon L. Thompson-Schill
Department of Psychology, University of Pennsylvania

Introduction

Learning which choices are good to make from the instruction of others is more time- and cost-efficient than learning via trial-and-error (incremental reinforcement learning). However, people will persist with choosing according to inaccurate instruction in the face of instruction-disconfirming feedback (instructional bias). The mechanisms underlying instructional bias are debated. Three competing models:

1. **Learning override**: Learning is unaffected, but overridden at action selection.
2. **Learning bias**: Learning is bolstered for instruction-congruent outcomes and suppressed for instruction-incongruent outcomes. Learning is unaffected, but overridden at action selection.
3. **Learning suppression**: Learning is suppressed for the instructed item.

Prior work showed that neural activity in early visual cortex (EVC) measured during a perceptual judgment task was sensitive to object value after learning. Based on this, we aim to use pattern analysis of EVC activity between intervals of value learning to track learning in the brain, independent of choice, and distinguish between these models of instructional bias.

Design

**Stimulus**

<table>
<thead>
<tr>
<th>Value (proportion of trials on which shape is correct)</th>
<th>A</th>
<th>B</th>
<th>C</th>
<th>D</th>
<th>E</th>
<th>F</th>
</tr>
</thead>
<tbody>
<tr>
<td>Training Pair</td>
<td>AB</td>
<td>CD</td>
<td>EF</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
| Training 75 trials per block. AB, CD, EF pairs.

Task: Choose the shape most likely to be correct. Choose feedback.

| Test 45 trials per block. All possible pairs of the 6 stimuli.
| Task: Choose the shape most likely to be correct. No choice feedback.

**Size Judgement**

6 trials per sub-block; 2 sub-blocks per stimulus per block

Task: Press a button when the shape size changes relative to the previous trial.

Planned fMRI Analyses: Pattern-similarity

Does neural activity in EVC reflect unbiased or biased object values?

**Size Judgement Block**

1 2 3 4 5

Neural Dissimilarity Matrix

Unbiased Learning Dissimilarity Matrix
(calculated from stimulus values)

Neural Dissimilarity Matrix calculated on each size judgement block is correlated with unbiased and biased learning dissimilarity matrices.

Biased Learning Dissimilarity Matrix
(calculated from stimulus values, D stimulus value is high)

Hypothesised fMRI results

Learning Override

1. Neural activity in EVC reflects unbiased stimulus values.
2. Classification accuracy for neural activity while viewing instructed and uninstructed stimuli increases at the same rate.

Learning Bias

1. Neural activity in EVC initially reflects biased stimulus values.
2. Classification accuracy for neural activity while viewing the instructed stimulus increases.

Learning Suppression

1. Neural activity in EVC reflect unbiased stimulus values.
2. Classification accuracy for neural activity while viewing the instructed stimulus increases at a slower rate.

References

1. Tardiff, Graves, & Thompson-Schill (2016). Frontiers in Human Neuroscience, 12, 472.

Acknowledgements

This research was funded by NIH R01 DC009209 awarded to Sharon Thompson-Schill.

Contact

Email: eprentis@sas.upenn.edu
Twitter: @EuanPrentis
Website: https://web.sas.upenn.edu/schill-lab/