Neural Correlates of Context Function in Associative Learning

Samuel Gershman (gershman@fas.harvard.edu)

Momchil Tomov (mtomov@g.harvard.edu)

Hayley Dorfman (hdorfman@g.harvard.edu)

Department of Psychology and Center for Brain Science,
Harvard University, 52 Oxford St., room 295.05
Cambridge, MA 02138, USA

Keywords: Bayesian modeling; Animal and human associative learning; Context-dependent learning

Introduction

Materials and Methods

Subjects. We recruited 28 healthy subjects (X female; Y-Z years of age; mean age Z +- SEM) to participate in this study. Eight participants were excluded due to excessive motion or insufficient data. The study was approved by the Harvard Institutional Review Board (IRB).

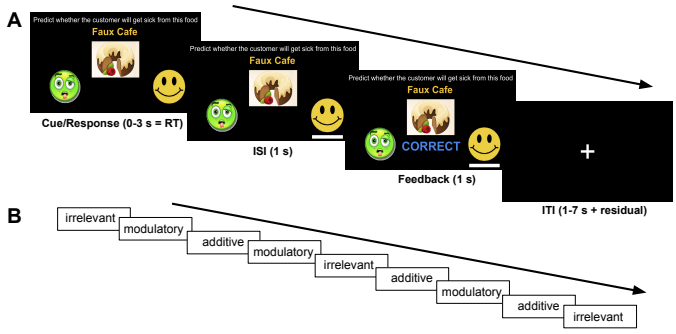

Figure 1: Experimental design. (A) Example timeline of events during a training trial. (B) Example sequence of blocks with the corresponding condition for each block.

Experimental design. We adapted the task used in Gershman (2017) with a mixed within-subjects design consisting of 9 blocks. Each block consisted of 20 training trials followed by 4 test trials. On each training trial, participants were asked to predict whether a particular food (the cue) in a particular restaurant (the context) would cause sickness (the outcome) and were subsequently informed whether their prediction was correct (Figure 1A). On each test trial, participants were asked to make a prediction about an old or a novel cue in an old or a novel context, without receiving any feedback, with each of the 2x2 combinations appearing exactly once. In each block, the cue-outcome contingencies depended on the context in accordance with one of the three causal interpretations which we refer to as the condition for that block. The nine blocks were divided in three consecutive groups such that each condition appeared in exactly one block in each group (Figure 1B). Thus each participant learned under each condition three times. Each block contained a different set of foods and restaurants that were randomized across blocks.

Simulations. We implemented the model presented in Gershman (2017). The model had two free parameters: the variance σ_w^2 of the Gaussian prior from which the weights are assumed to be drawn; and the inverse temperature β used in the logistic transformation from predictive posterior expectation to choice probability. Intuitively, the former corresponds to the level of uncertainty in the initial estimate of the weights, while the latter reflects the exploration-exploitation tradeoff of the model choices. We fit these parameters using maximum likelihood parameter estimation based on behavioral data obtained from 10 different subjects who performed the same task outside the scanner during a pilot version of the study (data not shown). The fitted values were $\sigma_w^2 = 0.1249$ and $\beta = 2.0064$. All other parameters had the same values as described in Gershman (2017). Each block was simulated independently using the same set of parameters.

Results

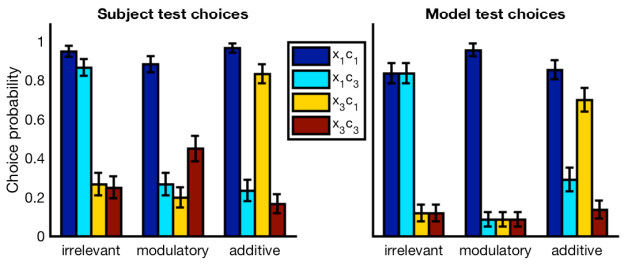

Figure 2: Subject (left) and model (right) performance on the test trials.

Behavioral performance. Test trial choices averaged across blocks are shown in Figure 2. Overall, participants learned the task quickly and exhibited the same within-subjects behavioral pattern as was previously reported using a between-subjects design (Gershman, 2017). The model successfully accounted for participants' choices on both the

training and the test trials ($r = 0.7283, p < 0.00001$) using the parameters obtained from pilot data from a different set of participants.

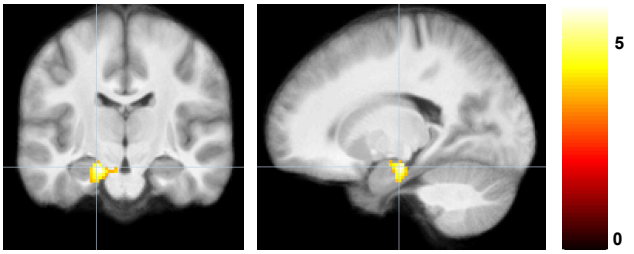

Figure 3: Transient activity related to the additive condition minus the irrelevant condition in left anterior hippocampus (MNI: -18 -16 -18, $t > 3.5518, p < 0.001$, cluster FWE corr). Right: T-value scale.

Imaging data. We had an a priori hypothesis that the hippocampus would be involved in modulating the cue-outcome association when it is influenced by the context. We therefore contrasted BOLD activation between blocks in the different conditions at the time of outcome presentation during the training phase. The BOLD signal did not differ significantly between the modulatory and the irrelevant conditions, nor between the modulatory and the additive conditions. The contrast between the additive and the irrelevant conditions showed increased activation in left anterior hippocampus (Figure 3; MNI coordinates of peak voxel: [-18 -16 -18]; t-value: 5.748; extent with $t > 3.5518$: 141; $p < 0.001$; cluster FWE corrected). Each contrast also included a constant regressor at the time of trial onset to account for the visual activation.

In order to measure the neural correlates of structure learning, we computed a contrast with a parametric modulator corresponding to the Kullback–Leibler divergence of the posterior over causal structures which is updated by the model at the time of outcome presentation on each training trial. The results are shown in Table 1 and Figure TODO. We found significant bilateral activation in the parietal cortex (the angular gyri), rostrolateral prefrontal cortex (rlPFC).

Table 1: Brain activation tracking the Kullback–Leibler divergence of the posterior over causal structures during training. Only cerebral regions with T-value > 5 are shown. All P-values are < 0.001 with cluster FWE correction. Regions were automatically labeled using the AAL2 atlas.

Brain region	Extent	T-value	MNI coordinates
Angular (R)	484	8.638	34 -64 48
Frontal Inf Oper (R)	341	8.378	48 20 34
Frontal Mid 2 (R)	130	7.440	36 56 -2
Frontal Mid 2 (L)	173	7.205	-42 56 2
Frontal Mid 2 (R)	86	6.996	34 12 54
Parietal Inf (L)	254	6.699	-30 -54 42
Parietal Sup (L)	254	5.566	-34 -72 54
Frontal Inf Tri (L)	173	6.583	-44 20 22
Temporal Inf (R)	15	6.461	60 -24 -20
Insula (L)	18	6.272	-28 22 -2
OFCant (R)	8	5.827	20 48 -16

General Formatting Instructions

The entire contribution of a short summary submission (including figures, references, and anything else) can be no longer than two pages. This short summary format is to be used for workshop and tutorial descriptions, symposia summaries, and publication-based presentation extended abstracts. Unlike submitted research papers, short summary submissions should *not* begin with a separate abstract. Prior to the first section of the short summary, there should be the header "**Keywords:**" followed by a list of descriptive keywords separated by semicolons, all in 9 point font, as shown above.

The text of the paper should be formatted in two columns with an overall width of 7 inches (17.8 cm) and length of 9.25 inches (23.5 cm), with 0.25 inches between the columns. Leave two line spaces between the last author listed and the text of the paper. The left margin should be 0.75 inches and the top margin should be 1 inch. **The right and bottom margins will depend on whether you use U.S. letter or A4 paper, so you must be sure to measure the width of the printed text.** Use 10 point Modern with 12 point vertical spacing, unless otherwise specified.

The title should be in 14 point, bold, and centered. The title should be formatted with initial caps (the first letter of content words capitalized and the rest lower case). Each author's name should appear on a separate line, 11 point bold, and centered, with the author's email address in parentheses. Under each author's name list the author's affiliation and postal address in ordinary 10 point type.

Indent the first line of each paragraph by 1/8 inch (except for the first paragraph of a new section). Do not add extra vertical space between paragraphs.

First Level Headings

First level headings should be in 12 point, initial caps, bold and centered. Leave one line space above the heading and 1/4 line space below the heading.

Second Level Headings

Second level headings should be 11 point, initial caps, bold, and flush left. Leave one line space above the heading and 1/4 line space below the heading.

Third Level Headings Third level headings should be 10 point, initial caps, bold, and flush left. Leave one line space above the heading, but no space after the heading.

Formalities, Footnotes, and Floats

Use standard APA citation format. Citations within the text should include the author's last name and year. If the authors' names are included in the sentence, place only the year

in parentheses, and Newell and Simon (1972), but otherwise place the entire reference in parentheses with the authors and year separated by a comma (Newell & Simon, 1972). List multiple references alphabetically and separate them by semicolons (Chalnick & Billman, 1988; Newell & Simon, 1972). Use the "et al." construction only after listing all the authors to a publication in an earlier reference and for citations with four or more authors.

Footnotes

Indicate footnotes with a number[1] in the text. Place the footnotes in 9 point type at the bottom of the column on which they appear. Precede the footnote block with a horizontal rule.[2]

Tables

Number tables consecutively. Place the table number and title (in 10 point) above the table with one line space above the caption and one line space below it, as in Table 2. You may float tables to the top or bottom of a column, or set wide tables across both columns.

Table 2: Sample table title.

Error type	Example
Take smaller	63 - 44 = 21
Always borrow	96 - 42 = 34
0 - N = N	70 - 47 = 37
0 - N = 0	70 - 47 = 30

Figures

Make sure that the artwork can be printed well (e.g. dark colors) and that the figures make understanding the paper easy. Number figures sequentially, placing the figure number and caption, in 10 point, after the figure with one line space above the caption and one line space below it, as in Figure **??**. If necessary, leave extra white space at the bottom of the page to avoid splitting the figure and figure caption. You may float figures to the top or bottom of a column, or set wide figures across both columns.

Acknowledgments

Place acknowledgments (including funding information) in a section at the end of the paper.

References Instructions

Follow the APA Publication Manual for citation format, both within the text and in the reference list, with the following exceptions: (a) do not cite the page numbers of any book, including chapters in edited volumes; (b) use the same format for unpublished references as for published ones. Alphabetize references by the surnames of the authors, with single author entries preceding multiple author entries. Order references by the same authors by the year of publication, with the earliest first.

Use a first level section heading, "**References**", as shown below. Use a hanging indent style, with the first line of the reference flush against the left margin and subsequent lines indented by 1/8 inch. Below are example references for a conference paper, book chapter, journal article, dissertation, book, technical report, and edited volume, respectively.

References

Chalnick, A., & Billman, D. (1988). Unsupervised learning of correlational structure. In *Proceedings of the tenth annual conference of the cognitive science society* (pp. 510–516). Hillsdale, NJ: Lawrence Erlbaum Associates.

Feigenbaum, E. A. (1963). The simulation of verbal learning behavior. In E. A. Feigenbaum & J. Feldman (Eds.), *Computers and thought.* New York: McGraw-Hill.

Gershman, S. J. (2017). Context-dependent learning and causal structure. *Psychonomic Bulletin and Review*, *24*, 557–565.

Hill, J. A. C. (1983). A computational model of language acquisition in the two-year old. *Cognition and Brain Theory*, *6*, 287–317.

Matlock, T. (2001). *How real is fictive motion?* Doctoral dissertation, Psychology Department, University of California, Santa Cruz.

Newell, A., & Simon, H. A. (1972). *Human problem solving*. Englewood Cliffs, NJ: Prentice-Hall.

Ohlsson, S., & Langley, P. (1985). *Identifying solution paths in cognitive diagnosis* (Tech. Rep. No. CMU-RI-TR-85-2). Pittsburgh, PA: Carnegie Mellon University, The Robotics Institute.

Shrager, J., & Langley, P. (Eds.). (1990). *Computational models of scientific discovery and theory formation*. San Mateo, CA: Morgan Kaufmann.

[1] Sample of the first footnote.

[2] Sample of the second footnote.