High-Level Perceptual Learning for Initially Ambiguous Stimuli

Logan T. Dowdle (logan.dowdle@maastrichtuniversity.nl) & Mario Senden (mario.senden@maastrichtuniversity.nl)
Department of Cognitive Neuroscience, Faculty of Psychology and Neuroscience, Maastricht University, P.O. Box 616
Maastricht, 6200 MD, Limburg, The Netherlands

Abstract

Perceptual skills are frequently studied using simple stimuli like Gabor patches to isolate basic perceptual mechanisms. However, in everyday expert decision-making, perceptual skills are deeply intertwined with higher-order cognition and semantic knowledge. For example, pathologists must accurately distinguish between various tissue types by interpreting complex visual patterns in microscopy images. To investigate how perceptual and semantic representations emerge for initially ambiguous stimuli and interact in the brain, we conducted a high-resolution 7T fMRI pilot study in which a lay participant learned to distinguish tissue with high and low tumor-infiltrating lymphocyte (TIL) count from histopathology images. The participant underwent daily perceptual training with feedback, paired with at-home semantic study of cancer types. fMRI data were collected in five sessions distributed over the training period wherein the participant performed a TIL classification task. This allowed us to track behavioral learning and corresponding changes in cortex-wide neural representations.

Keywords: visual processing; perceptual learning; decision making; fmri; semantic representations; cognition

Introduction

Perceptual skills are frequently studied using simple stimuli like Gabor patches to isolate basic perceptual mechanisms (Adini, Sagi, & Tsodyks, 2002; Schoups, Vogels, Qian, & Orban, 2001; Lange, Senden, Radermacher, & De Weerd, 2020). While this approach has yielded valuable insights into the foundational aspects of perception, it does not capture the complexity of perceptual skills as they operate in real-world contexts (Tsushima, Sawahata, & Komine, 2020). For example, a sommelier not only discriminates between wines based on subtle sensory cues but also recalls detailed semantic information about each wine's composition, origin, and ideal food pairings. Similarly, pathologists must accurately distinguish between various tissues under different disease conditions. This perceptual discrimination is closely linked with semantic knowledge about diseases, including typical symptoms, prevalence, and prognostic implications. Despite the critical importance of integrating perceptual and semantic information, it is currently not well established how low-level perceptual skills relate to semantic knowledge at the neural level.

We provide a pilot investigation into this issue by examining how perceptual expertise develops in a complex, semantically rich domain and how it is reflected in neural representations across the cortical hierarchy. Specifically, we trained a lay participant to discriminate histopathology images of cancerous tissue according to tumor infiltrating lymphocyte (TIL) density and simultaneously exposed them to semantic information about cancer types. Functional MRI data were collected at 7 Tesla while the participant performed a TIL classification task. A total of five scan sessions were distributed over the learning period. Using representational connectivity (RC), a measure of pairwise similarity between region-specific representational geometries, we show a progressive alignment of representations between early visual, high-level visual, attention, cognitive control and motor planning regions.

Methods

A single participant (Male, 39 years old) naive to pathology completed the experiment over a period of 4 weeks.

Stimuli

We used a subset of a publicly available collection of whole slide microscopy images containing high and low tumor-infiltrating lymphocyte (TIL) counts (Abousamra et al., 2022). We split the dataset into at-home training and in-scanner stimulus groups, uniformly drawn from 12 cancer types.

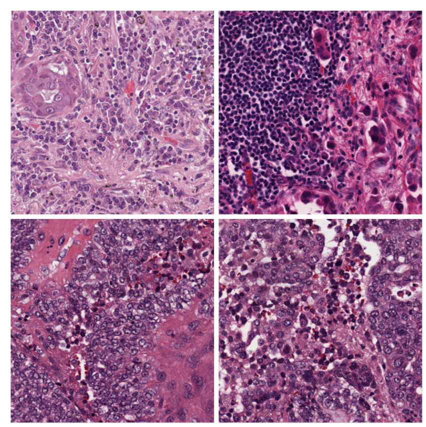

Figure 1: Example stimuli Left and right column show examples of low and high TIL counts, respectively. Top row shows lung and bottom row ovarian tissue.

Offline Training

At home training used 1,158 stimuli with PsychoPy (Peirce et al., 2019). This involved a repeated presentations of a single image, which remained on screen until the participant indicated a high or low TIL count. The script then provided feedback and reported cancer type. The participant completed 17 sessions and also used online resources to gain expertise.

fMRI Procedures and Processing

We performed 5 imaging sessions during a 4-week period. The first session occurred before any training. Each image was presented for 4.5s with a 1s blank during which the subject indicated high/low TIL via a button press (no feedback). Following an inter-stimulus interval (0 to 4.5 s, drawn from Poisson), the procedure repeated. 140 stimuli (3 repetitions) were shown per session, 120 repeated across all sessions and 20 unique to each session for a total of 420 per session.

In each session we collected 10 runs of whole-brain 7T BOLD data (1.6mm iso., TR:1.5s, TE:21.2ms, 84 slices) with 238 volumes and anatomical images.

Preprocessing used AFNI (Cox, 1996) and ANTS (Avants, Epstein, Grossman, & Gee, 2008) for slice-timing, distortion and motion correction and run/session/anatomical alignment. The anatomical data was processed with FreeSurfer's (Dale, Fischl, & Sereno, 1999) recon-all (version 8.0.0beta) to generate cortical surfaces and align to fsaverage. We used glmsingle (Prince et al., 2022) to estimate single trial betas. These betas were projected into fsaverage space and analyzed using the Glasser atlas (Glasser et al., 2016).

Results

Behavioral Results

To assess learning, we analyzed the participant's accuracy in discriminating between high and low TIL counts using a binomial GLM with correctness of trial-level responses as outcome and session, cancer (tissue) type and their interaction as predictors. Wald tests revealed a significant overall improvement in accuracy across the five sessions (main effect of session: Wald $\chi^2(1) = 10.15,\ p = 0.0014$). Importantly, the learning effect differed per cancer type with a significant interaction between session and cancer type (Wald $\chi^2(11) = 32.37$, $p = 0.0007$). Follow-up simple effects analyses per cancer type indicated that this learning effect was driven by significant improvements in classification accuracy over sessions specifically for liver ($z = 4.9,\ p < 0.001$), prostate ($z = 2.2$, $p = 0.03$) and breast ($z = 3.2,\ p = 0.001$) tissue.

Representational Connectivity

To assess the progression of representational geometries across the cortex and time, we computed RC in each session. To that end, we first obtained representational dissimilarity matrices (RDMs) within each ROI for the 120 stimuli shown in every session using a correlation-based distance metric. Representational connectivity is then the similarity between RDMs of every pair of ROIs, here quantified using Spearman's ρ. We performed a regression to assess whether connections exhibit a linear trend over sessions for each pair of ROIs and established significance of linear trends via a bootstrapping procedure with 50,0000 samples. A total of 126 connections involving 115 ROIs were significant after Bonferroni correction. Session 5 RC among the 115 identified ROIs as well as difference between session 5 and session 1 RC are shown in Figure 2.

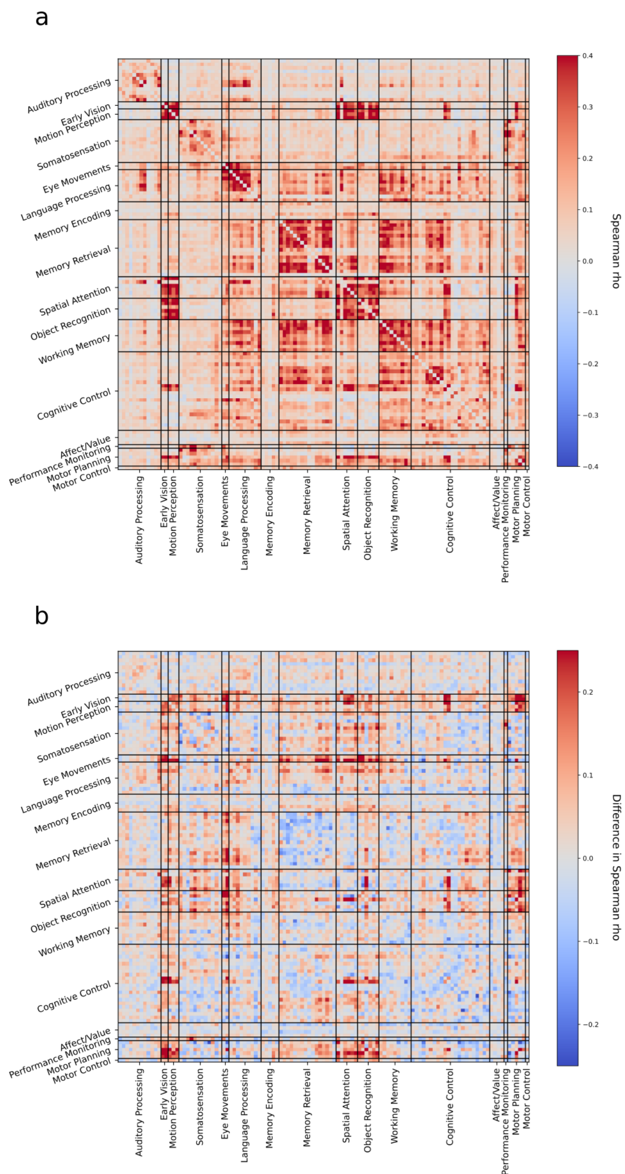

Figure 2: Representational Connectivity. **a** RC in session 5 limited to ROIs that exhibit significant connectivity changes across sessions. ROIs are arranged by functional domain. ROI labels not shown. **b** Difference in RC between sessions 5 and 1. ROIs and their arrangement as in panel a.

Discussion

Our results provide a first indication that people can learn new semantic meanings from *ambiguous* stimuli. We found behavioral evidence of this effect in a lab paradigm. Most importantly, this is accompanied by progressive alignment of representations between early visual, high-level visual, attention, cognitive control and motor planning regions.

References

Abousamra, S., Gupta, R., Hou, L., Batiste, R., Zhao, T., Shankar, A., ... others (2022). Deep learning-based mapping of tumor infiltrating lymphocytes in whole slide images of 23 types of cancer. *Frontiers in oncology*, *11*, 806603.

Adini, Y., Sagi, D., & Tsodyks, M. (2002, February). Context-enabled learning in the human visual system. *Nature*, *415*(6873), 790–793. Retrieved from `https://doi.org/10.1038/415790a` doi: 10.1038/415790a

Avants, B. B., Epstein, C. L., Grossman, M., & Gee, J. C. (2008, February). Symmetric diffeomorphic image registration with cross-correlation: Evaluating automated labeling of elderly and neurodegenerative brain. *Medical Image Analysis*, *12*(1), 26–41. doi: 10.1016/j.media.2007.06.004

Cox, R. W. (1996, June). AFNI: Software for Analysis and Visualization of Functional Magnetic Resonance Neuroimages. *Computers and Biomedical Research*, *29*(3), 162–173. doi: 10.1006/cbmr.1996.0014

Dale, A. M., Fischl, B., & Sereno, M. I. (1999, February). Cortical Surface-Based Analysis: I. Segmentation and Surface Reconstruction. *NeuroImage*, *9*(2), 179–194. doi: 10.1006/nimg.1998.0395

Glasser, M. F., Coalson, T. S., Robinson, E. C., Hacker, C. D., Harwell, J., Yacoub, E., ... Van Essen, D. C. (2016, August). A multi-modal parcellation of human cerebral cortex. *Nature*, *536*(7615), 171–178. doi: 10.1038/nature18933

Lange, G., Senden, M., Radermacher, A., & De Weerd, P. (2020). Interfering with a memory without erasing its trace. *Neural Networks*, *121*, 339–355.

Peirce, J., Gray, J. R., Simpson, S., MacAskill, M., Höchenberger, R., Sogo, H., ... Lindeløv, J. K. (2019, February). PsychoPy2: Experiments in behavior made easy. *Behavior Research Methods*, *51*(1), 195–203. doi: 10.3758/s13428-018-01193-y

Prince, J. S., Charest, I., Kurzawski, J. W., Pyles, J. A., Tarr, M. J., & Kay, K. N. (2022, November). Improving the accuracy of single-trial fMRI response estimates using GLMsingle. *eLife*, *11*, e77599. doi: 10.7554/eLife.77599

Schoups, A., Vogels, R., Qian, N., & Orban, G. (2001). Practising orientation identification improves orientation coding in v1 neurons. *Nature*, *412*(6846), 549–553.

Tsushima, Y., Sawahata, Y., & Komine, K. (2020). Task-dependent fmri decoder with the power to extend gabor patch results to natural images. *Scientific Reports*, *10*. doi: 10.1038/s41598-020-58241-x