Brain and Cognition 116 (2017) 63–70

Contents lists available at ScienceDirect

Brain and Cognition

journal homepage: www.elsevier.com/locate/b&c

The effect of integration masking on visual processing in perceptual categorization

Sébastien Hélie

Department of Psychological Sciences, Purdue University, United States

ARTICLE INFO

Keywords:
Categorization
Integration masking
Visual processing
fMRI
MVPA

ABSTRACT

Learning to recognize and categorize objects is an essential cognitive skill allowing animals to function in the world. However, animals rarely have access to a canonical view of an object in an uncluttered environment. Hence, it is essential to study categorization under noisy, degraded conditions. In this article, we explore how the brain processes categorization stimuli in low signal–to–noise conditions using multivariate pattern analysis. We used an integration masking paradigm with mask opacity of 50%, 60%, and 70% inside a magnetic resonance imaging scanner. The results show that mask opacity affects blood–oxygen–level dependent (BOLD) signal in visual processing areas (V1, V2, V3, and V4) but does not affect the BOLD signal in brain areas traditionally associated with categorization (prefrontal cortex, striatum, hippocampus). This suggests that when a stimulus is difficult to extract from its background (e.g., low signal–to–noise ratio), the visual system extracts the stimulus and that activity in areas typically associated with categorization are not affected by the difficulty level of the visual conditions. We conclude with implications of this result for research on visual attention, categorization, and the integration of these fields.

1. Introduction

Learning to recognize and categorize objects is an essential cognitive skill allowing animals to function in the world (Ashby, 2013). For example, recognizing another animal as a friend or a foe allows for determining how to interact with it. Likewise, recognizing a plant as edible (or not) can ensure survival. However, animals rarely have access to a canonical view of an object in an uncluttered environment (Hegdé, Thompson, Brady, & Kersten, 2012). The same objects are often seen with a different viewpoint, partially obstructed, or in less than ideal lighting conditions. Hence, it is essential to study categorization under noisy, degraded conditions.

Towards this end, Hélie and Cousineau (2015) recently studied the effect of backward masking and integration masking on human category learning. The results show that reducing the delay between the stimulus presentation and the mask (backward masking) reduces categorization accuracy, and that the reduction in accuracy is more important for non–verbal (information–integration) categorization then for verbal (rule–based) categorization. A second experiment shows that increasing the mask opacity when the stimulus and mask are presented simultaneously at the same location (integration masking) has the same effect as reducing the delay between the mask and stimulus. Specifically, increasing the opacity of the mask reduces categorization accuracy, and the reduction in accuracy is more important for non–verbal categorization than for verbal categorization. Hélie and Cousineau (2015) argued that both backward masking and integration masking affect the signal–to–noise ratio (SNR), and that non–verbal categorization relies on mental representations that are less robust to noise than the mental representations supporting verbal categorization.

One follow–up question to the Hélie and Cousineau (2015) experiments is how does the brain process the categorization stimuli in low SNR conditions? One possibility is that areas typically associated with visual processing in posterior cortex (e.g., V1, V2, V3, V4; Roe et al., 2012) extract the stimulus from background noise, and that areas typically associated with categorization [e.g., striatum, prefontal cortex (PFC), hippocampus (HC); Hélie, Roeder, and Ashby, 2010; Seger and Miller, 2010] are not affected by the SNR. Another possibility is that visual processing is similar with low and high SNR, but that the categorization system received a degraded stimulus representation in low SNR conditions and needs to adjust its processing accordingly.

To disentangle these possible explanations, we replicated the verbal integration masking condition of Hélie and Cousineau (2015) inside a Magnetic Resonance Imaging (MRI) scanner. Mask opacity of 50%, 60%, and 70% were used because they have been shown to yield similar categorization accuracy. We hypothesize that mask opacity should affect brain areas related to visual processing (e.g., V1 – V4) but not brain areas related to category learning (e.g., striatum, PFC, HC). This hypothesis is based on the following: (1) during a recognition test, the left

E-mail address: shelie@purdue.edu.

http://dx.doi.org/10.1016/j.bandc.2017.06.001
Received 2 December 2016; Received in revised form 17 May 2017; Accepted 1 June 2017
Available online 09 June 2017

fusiform gyrus is differently activated depending on the amount of clutter in which each stimulus was presented during the learning phase (Hegdé et al., 2012); (2) learning increases the amount of information communicated by neurons in V4 for degraded objects in image recognition by rhesus monkeys (Rainer, Lee, & Logothetis, 2004); and (3) neurons in the PFC fire in a similar manner to degraded and non–-degraded stimuli after training rhesus monkeys in a delayed matching–to–sample task (Rainer & Miller, 2000).

Results (1) and (2) suggest differential processing in visual areas related to difficulty in extracting the stimulus from background information. Specifically, Hegdé et al. (2012), manipulated the clutter, which can make the stimulus harder to isolate and extract (1). This added difficulty is also present in integration masking. Likewise, Rainer et al. (2004) showed that V4 neurons learn to increase the amount of information communicated for degraded stimuli, which suggests extracted or de–noised information (2). This result likely generalizes to integration masking because the mask degrades the stimulus. Finally, Rainer and Miller (2000) suggest an absence of the effect of stimulus degradation on a brain area important for categorization, namely the PFC (3). For the same reason as Rainer et al. (2004), this result should also generalize to integration masking. This study expends on the Rainer and Miller (2000) study in that it uses functional MRI (fMRI) to look at whole–brain activity. To anticipate, the results support the hypothesis and show that mask opacity affects blood–oxygen–level dependent (BOLD) signal in visual processing areas (V1, V2, V3, and V4) but does not affect the BOLD signal in brain areas traditionally associated with categorization (PFC, striatum, HC).

2. Material and methods

The experiment used the same stimuli, masks, and categories as the rule–based condition in Hélie and Cousineau (2015) Experiment 2. The main differences were (1) the experiment lasted 2 sessions, with the second session of the experiment conducted in an MRI scanner, (2) the timing of the events was jittered, (3) only mask opacity of 50%, 60%, and 70% were used and, (4) the experiment design was within–subject.

2.1. Participants

Twenty students at Purdue University were recruited to participate in the experiment (9 males, 11 females). All participants gave their written informed consent to participate in the study. The institutional review board of Purdue University approved all procedures in this experiment. All the participants received a monetary compensation of $50 to participate in the experiment. One male participant dropped out of the experiment due to claustrophobia in the scanner, so the final sample included 19 participants.

2.2. Stimuli and apparatus

The stimuli were circular sine–wave gratings of constant contrast and size. An example stimulus is shown in Fig. 1a. Each stimulus was defined by a set of points (x_1, x_2) sampled from an arbitrary 100×100 stimulus space and converted to a disk using the following equations: frequency (bar width) $= \frac{x_1}{30} + 0.25$ cycles per degree (cpd), and bar orientation $= \frac{9x_2}{10} + 20$ degrees. This yielded stimuli that varied in orientation from 20° to 110° (counterclockwise from horizontal) and in bar width (frequency) between 0.25 and 3.58 cpd. The stimuli were generated with Matlab using the Psychophysics Toolbox (Brainard, 1997). In each trial, a single stimulus occupying about 5° of visual angle was presented in the center of the display. The stimuli were separated in two categories and were generated using the randomization technique of Ashby and Gott (1988). Fig. 1b shows the stimulus categories. Category "A" stimuli were generated from two multivariate normal distributions with the following parameters: $\mu_{A1} = (30,50); \Sigma_{A1} = \begin{pmatrix} 10 & 0 \\ 0 & 150 \end{pmatrix}$ and $\mu_{A2} = (50,70); \Sigma_{A2} = \begin{pmatrix} 150 & 0 \\ 0 & 10 \end{pmatrix}$. A similar sampling method was used to generate category "B" stimuli: $\mu_{B1} = (50,30); \mu_{B2} = (70,50); \Sigma_{B1} = \Sigma_{A1}$; and $\Sigma_{B2} = \Sigma_{A2}$. Note that each trial within a block showed a unique stimulus, and that perfect accuracy was possible using a simple verbal rule: "If the bars are thin and the orientation is near horizontal, press B; Otherwise, press A".

In each trial, a mask covered the stimulus to produce integration masking (Breitmeyer & Ogmen, 2006). The masks were taken from Hélie and Cousineau (2015). Briefly, each mask was uniquely generated using a random cloud of dots in Fourier space and transformed into a visual mask using an inverse Fourier transform. The details of mask generation can be found in Hélie and Cousineau (2015). An example mask is shown in Fig. 1c. Each mask was centered in the screen and had three times the height and the width of the stimulus. Stimulus presentation, mask presentation, feedback, and response recording were controlled and acquired using Matlab. During the whole experiment, the screen background was gray.

The experiment consisted of two training sessions on consecutive workdays. The first session was performed outside the scanner on a regular desktop computer while the second session was conducted inside an MRI scanner. During the first session, the stimuli and masks were presented on a 21–inch monitor (1280×1024 resolution). Responses were given on a standard keyboard: the "d" key for an "A" response and the "k" key for a "B" response (identified with stickers labeled "A" and "B" respectively). During the scanning session, the stimuli and masks were presented using NordicNeuroLab goggles with a screen resolution of 800×600. The participants selected category "A" or "B" responses using two button boxes (one in each hand) by Current Designs Inc. Consistent with the response key assignment in the first session, the button box in the left hand indicated an "A" category response and the button box in the right hand indicated a "B" category response.

2.3. Study design

Participants were told that they were taking part in a categorization experiment and that they were to assign each stimulus into either an "A" or "B" category. The number of blocks, trials, and the timing of each trial was identical in both sessions. Each session was composed of 6 blocks of 78 trials. In each block, half the stimuli were members of category "A" while the other half were members of category "B". One third of the "A" stimuli were covered by a mask with an opacity of 50%, another third was covered by a mask with an opacity of 60%, and the last third was covered by a mask with an opacity of 70%. The same applies to category "B" stimuli. Fig. 1(d)–(f) show example stimuli covered by masks with opacities of 50–70% (respectively). These mask opacity levels were selected because they have been shown to produce similar categorization accuracy with the categories shown in Fig. 1b (Hélie & Cousineau, 2015). Each stimulus and each mask within a given block is unique. Each block contains a (reshuffled) copy of the same 78 stimuli and masks, which allows for different ordering and stimulus/mask pairing.

The timing of a trial (for both sessions) is shown in Fig. 1g. Each stimulus was presented for 2000 ms. Correct responses were followed by a green check mark displayed for 2000 ms. Incorrect responses were followed by a red "X" mark displayed for 2000 ms. If participants did not respond before the stimulus disappeared, a black dot was displayed for 2000 ms. A fixation crosshair appeared for 1000 ms before the stimulus on an average of 48% of the trials (see caption of Fig. 1 for details). The crosshair was used to focus the participant's attention before stimulus presentation, which is standard in perceptual categorization experiments. Importantly, the crosshair was not diagnostic of the stimulus category membership and the participants had more than enough time to make a categorization decision with or without the crosshair. However, the irregular presentation of the crosshair in the