The time course of novel visual object recognition

Martin Arguin & Justine Massé
Centre de Recherche en Neuropsychologie et Cognition
Département de psychologie, Université de Montréal, Montréal, Canada.

Visit us at https://www.neurocognitionvisionlab.org/

INTRODUCTION

While there is increasing physiological (e.g. ERPs and spectral analysis of EEG) and behavioural evidence (e.g. Dugué et al., 2018 VSS Abstracts) that the acquisition of new information by the human visual system is temporally discontinuous, our knowledge in this regard remains very limited.

Our lab has developed a novel behavioural method to determine the temporal features of stimulation which impact visual processing effectiveness (see Refs below). Here, the method is applied in a task of novel object recognition. The discussion will offer an overview of the stimulus class effects on the temporal features which affect the effectiveness of visual processing.

METHODS

Participants: 12 adults with normal or corrected vision.

Training phase: Participants were first trained to associate each of six novel objects (Fig. 1) displayed in variable 3D orientations to a particular digit (from 1 to 6). Training was then pursued in an identification task with unaltered objects presented individually. The experimental test begun only when participants were over 90% accurate in a block of 192 practice trials.

Experimental phase: Object identification task with stimuli made from the additive combination of signal (the target object) and noise (white noise). Signal/noise ratio (SNR) varied randomly throughout the 200 ms stimulus exposure which occurred at a 120 Hz rate. Each participant completed a total of 1200 trials.

NOVEL OBJECTS

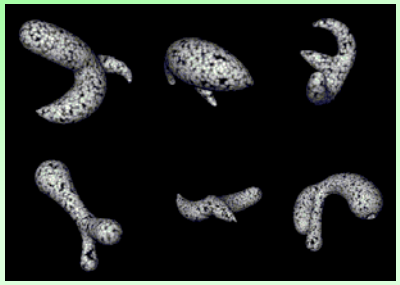


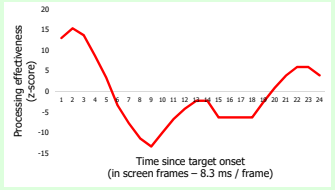
Figure 1. Collection of novel objects used in the Experiment. In both the training and experimental phases, objects were displayed in one of four widely disparate 3D orientations.
Objects were generated using a program graciously provided by Prof. C. E. Connor, School of Medicine, Johns Hopkins University.

Data analyses: The temporal properties of the stimulation on each trial were coded either as the raw SNR, the Fourier features (amplitude and phase of each constituent frequency) of SNR, or as the outcome of a time-frequency analysis of SNR.

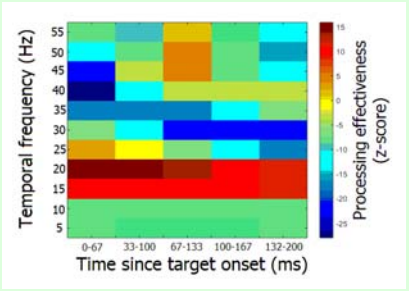
Classification images for each participant were constructed from the weighted subtraction of stimulus properties associated to errors from those associated to correct responses. These classification images representing processing effectiveness were transformed into Z scores by bootstrapping (1000 iterations) and then averaged. The average classification images were smoothed and then submitted to the Pixel test (Chauvin et al., 2005) to reveal significant differences. All the distinct consecutive values in the classification images that follow are significantly different from one another ($p < .05$).

RESULTS

Raw SNR classification image



Time-frequency classification image



Fourier Phase x Amplitude classification image: The distribution of Fourier amplitudes and phases (binned into 12 distinct levels) for the temporal frequencies included in the SNR (from 5 to 55 Hz in 5 Hz steps) was totally uniform (i.e. no significant difference) and is therefore not illustrated.

Between-subject consistency: The average of all possible pairwise correlations between individual classification images (between-subject consistency) was about null (0.03) for the raw SNR classification images. In contrast, it was quite high (0.65) for the time-frequency classification images.

DISCUSSION

Visual processing effectiveness in a task of novel object recognition is discontinuous through the first 200 ms of stimulus exposure and appears to involve set of oscillatory mechanisms in the range of 5 to 55 Hz. The high between-subject consistency of the time-frequency classification images suggests that the temporal stimulus properties it captures are fundamental for the human visual system.

Using the exact same methods as here, we have conducted other experiments investigating the visual recognition of words, familiar objects and faces.

The raw SNR and time-frequency classification images reveal highly distinct patterns of temporal features characterizing the processing of each stimululs class.

Between-subject consistency of the time-frequency classification images is remarkably high for highly optimized visual recognition tasks (with words: 0.87, and familiar objects: 0.93) and notably weaker for less practiced tasks (non-familiar objects and faces of famous actors: 0.67).

REFERENCES

Achouline et al. (2019). VSS Abstracts. || Arguin (2018). VSS Abstracts. || Blais et al. (2013). *Cognition*, **128**, 353-362. || Chauvin et al. (2005). *Journal of Vision*, **5**, 659-667. || Dugué et al. (2018). VSS Abstracts. || Ferrandez et al. (2019). VSS Abstracts. || Fortier-St-Pierre et al. (2019). VSS Abstracts.

Acknowledgements: Supported by a grant from the Natural Sciences and Engineering Research Council of Canada and the Fonds de Recherche Québec – Nature et Technologies.