coneural
Center for Cognitive and Neural Studies

Visual Sampling and Integration of Information in Object Recognition

Ioana Ţincaş, Vasile V. Moca and Raul C. Mureşan

Center for Cognitive and Neural Studies
Romanian Institute of Science and Technology
Cluj-Napoca, Romania

✉ tincas@coneural.org

Abstract

Eye movement control is relatively well understood in reading, but less so in object recognition, where research has been focused mainly on perception of scenes containing multiple objects. The current study aims to investigate eye movements during visual recognition of individual objects. The main obstacle in achieving this goal is the fact that object identification tends to be extremely fast (usually within the time-span of a single fixation). To prolong this quasi-instantaneous process and force participants to sample and integrate visual information across multiple fixations, we applied the "Dots" method developed in our laboratory (Moca et al., 2011). Starting from a source image, this method identifies regions containing contour information and then deforms a lattice of dots to represent these regions in a controlled fashion. The resulting stimulus can contain an arbitrarily small amount of information about the original image, thus being more difficult to recognize. Here we used photographic source images representing either coherent or scrambled objects. Ten healthy young adults were asked to discriminate between these two categories, and to correctly name the coherent objects. Results indicate that our method was successful in inducing participants to generate a relatively high number of fixations before reaching a decision. Additionally, exploration patterns were different for the two categories of stimuli: when viewing coherent objects, participants generated a lower number of (longer) fixations, and had a tendency to sample and integrate less of the lattice deformation, but a similar amount of underlying contour information.

INTRODUCTION

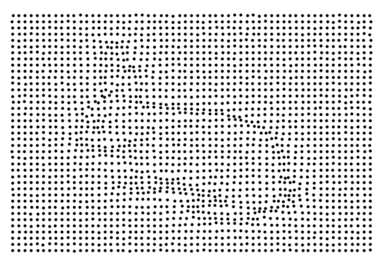
- Eye movement control (EMC) refers to factors affecting *where*, *when* and *how* our eyes move during visual exploration of the surrounding environment. Previous research shows that these aspects are affected by factors such as stimulus quality (Henderson, 2003; Henderson et al., 2013) or task set (Castelhano et al., 2009; Mills et al., 2011).
- EMC has been investigated mainly using visual search or memory tasks performed on visual scenes (see e.g., Castelhano et al., 2009; Mills et al., 2011).

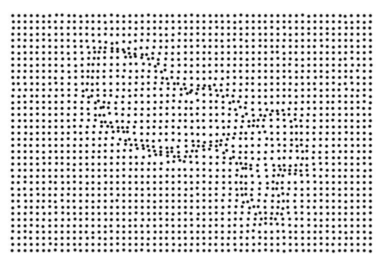
- **Questions:**
  - **How do people explore visual information when the task is object identification (recognition)?**
  - **Is this exploration pattern "controlled" by the informational content of the stimulus?**

METHOD

- **Participants**
  - 9 adults (6 females), aged 19-32. (Ten participants were tested initially; one was eliminated due to overly noisy eye tracking data).

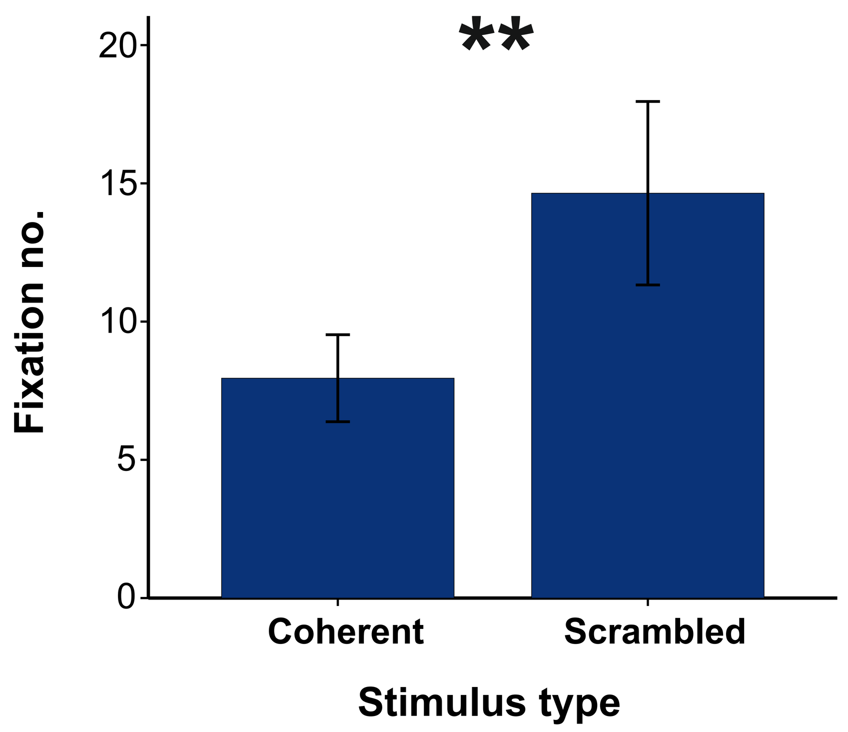
- **Stimuli, Task, Procedure**
  - 80 stimuli were generated using a procedure developed in our lab (details: Moca et al., 2011). Size: 8.7° x 5.6°; viewing distance: 1.12 m
  - Two types of stimuli: *Coherent* (meaningful) vs *Scrambled* (meaningless).
  - Task: decide whether the stimulus is Coherent (and identify) or Scrambled.
  - No time limit
  - Eye tracking: ASL EyeStart 6000; 50 Hz


Coherent

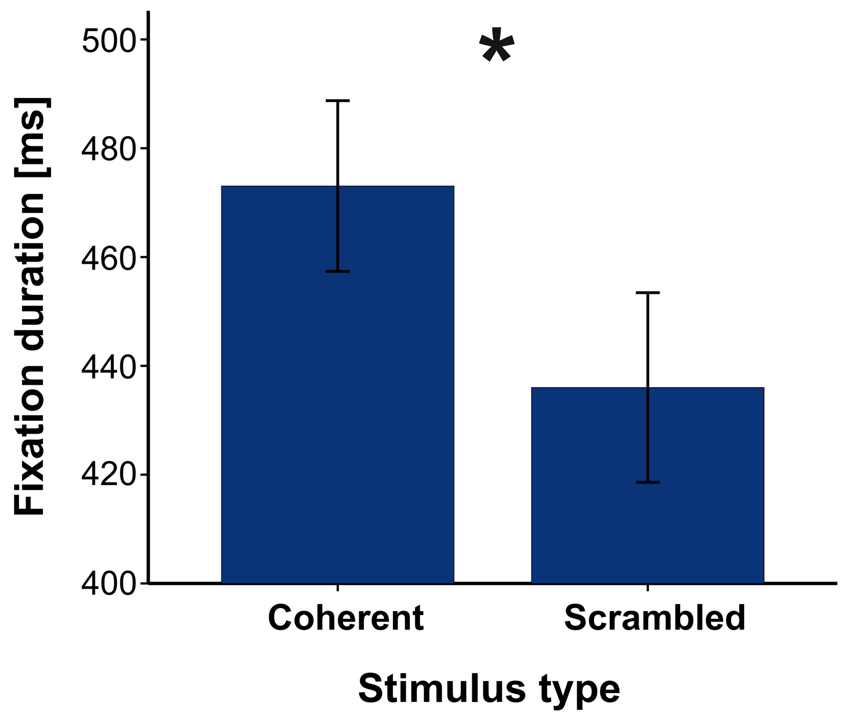

Scrambled

RESULTS

Fixation statistics



Wilcoxon's $Z = 2.67$, $p < .01$

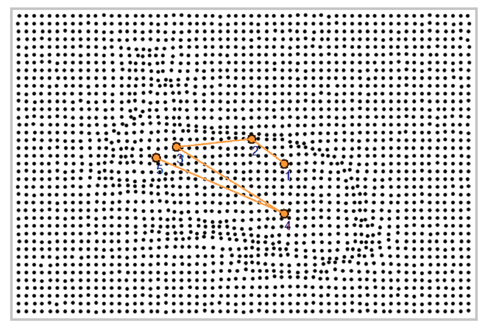


Wilcoxon's $Z = 2.19$, $p < .05$

Accuracy rate: 81.25 - 98.75 % correct
Trial duration: $Med_{Coherent}$ = 3181.60; $Med_{Scrambled}$ = 4777.07

Visual sampling and integration

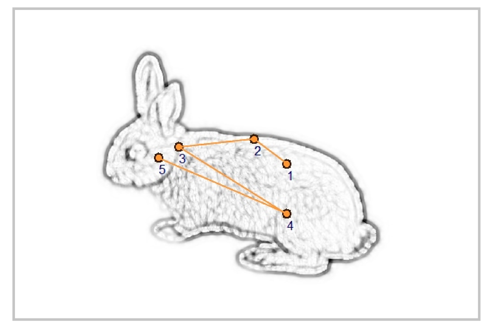
Dot Displacement (DD)



Amount of dot displacement (in pixels) from original position on the grid.

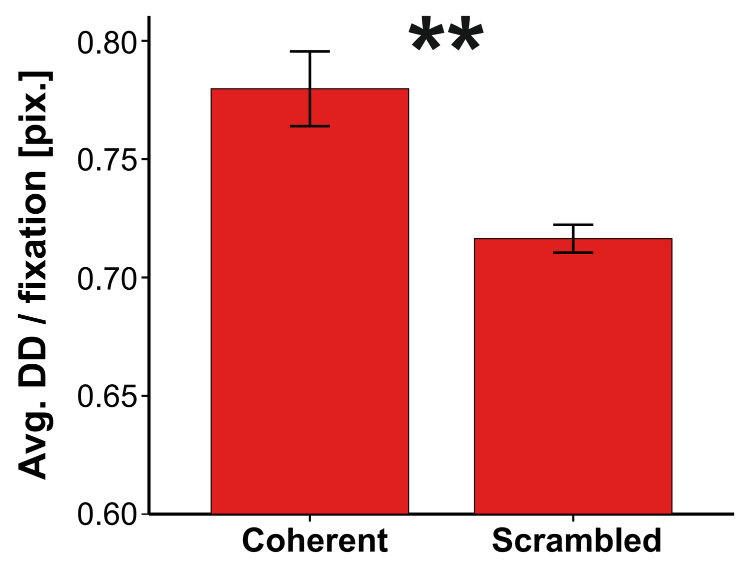
0.5° around each fixation

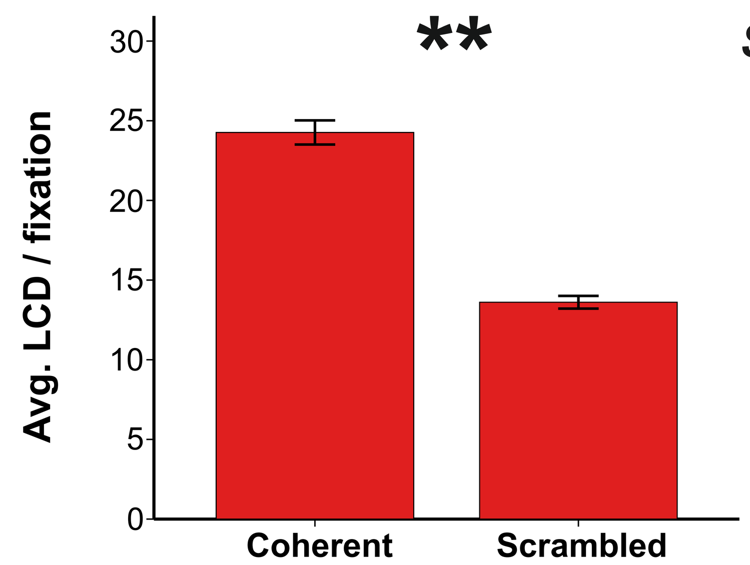
Local Contour Density (LCD)



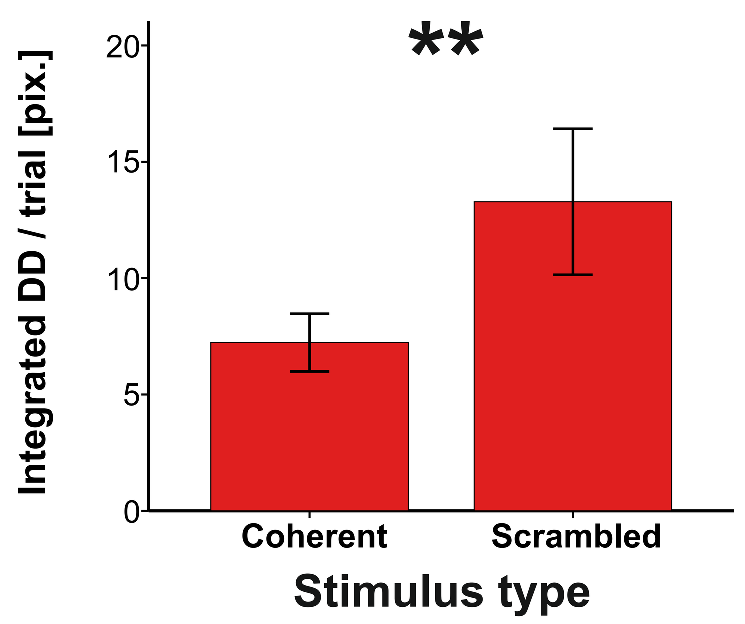
Based on the contrast map of the starting image.

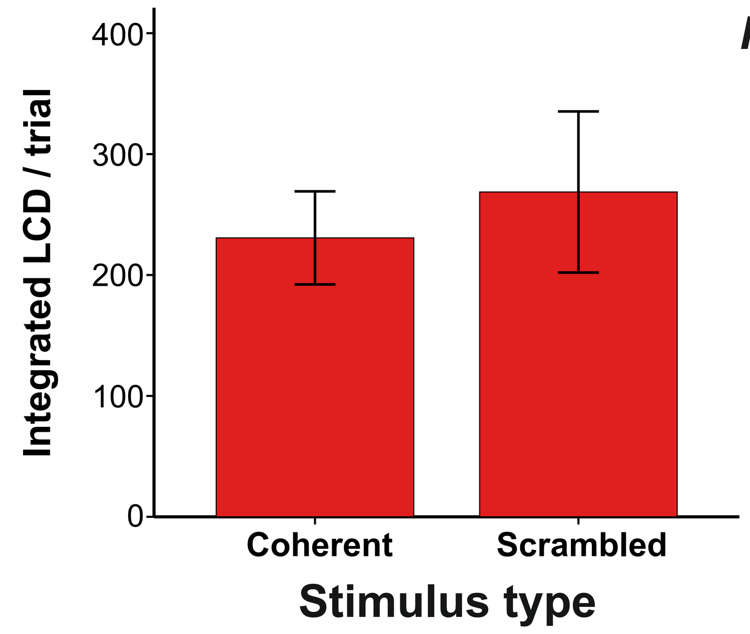
0.5° around each fixation





Sampled information = average amount of DD/LCD at each fixation location





Integrated information = amount of DD/LCD summed up over all fixation locations within a trial

**Wilcoxon's $Z = 2.67$, $p < .01$
Wilcoxon's $Z = 1.01$, *ns*

CONCLUSIONS

- **Coherent (vs. Scrambled) stimuli are associated with *fewer* (but *longer*) fixations.** Partly congruent with previous research: increased task difficulty has been linked to more and longer fixations (Castelhano et al., 2008; Henderson et al., 2013; Mills et al., 2011).

- When viewing Coherent (vs. Scrambled) stimuli, (1) **fixations tend to "land" on more informative areas**, and (2) participants need to integrate **less of the dot displacement**, but **a similar amount of contour information** to reach a correct decision.

- Automatic saccade programming (i.e., the decision of *when* to move the eyes to a new location) can be modulated by the difficulty of processing the current visual input (Henderson et al., 2013; Nuthman et al., 2010), as well as fixation history (Hooge et al., 2007).

- Our results suggest that the decision to terminate a fixation might also be influenced by the *amount of relevant information* contained in the currently fixated area. (*Relevance* = congruency between current input and the already-integrated information). Highly relevant visual input would delay the next saccade until the current sensory information has been processed and integrated.

References

Castelhano, M. S., Mack, M. L., & Henderson, J. M. (2009). Viewing task influences eye movement control during active scene perception. *Journal of Vision, 9*(3).

Henderson, J. M. (2003). Human gaze control during real-world scene perception. *Trends in Cognitive Sciences, 7*(11), 498–504.

Henderson, J. M., Nuthmann, A., & Luke, S. G. (2013). Eye movement control during scene viewing: Immediate effects of scene luminance on fixation durations. *Journal of Experimental Psychology: Human Perception and Performance, 39*(2), 318–322.

Hooge, I. T. C., Vlaskamp, B. N. S., & Overi, E. A. B. (2007). Saccadic search: On the duration of a fixation. In R.P.G. Van Gompel et al. (Eds), *Eye movements: A window on mind and brain* (pp. 581-596). Oxford, UK: Elsevier.

Mills, M., Hollingworth, A., Stigchel, S. V. der, Hoffman, L., & Dodd, M. D. (2011). Examining the influence of task set on eye movements and fixations. *Journal of Vision, 11*(8).

Moca, V. V., Ţincaş, I., Melloni, L., & Mureşan, R. C. (2011). Visual exploration and object recognition by lattice deformation. *PLoS ONE, 6*(7), e22831.

Nuthmann, A., Smith, T. J., Engbert, R., & Henderson, J. M. (2010). CRISP: A computational model of fixation durations in scene viewing. *Psychological Review, 117*, 382-405.

Acknowledgements

This work was supported by a grant of the Romanian National Authority for Scientific Research, CNCS – UEFISCDI, project no. PN-II-RU-PD-2011-3-0278.
We are grateful to the volunteers who agreed to participate in this study.