Coding of Visual Stimuli for Size and Animacy

Xiaokun Xu[1*], Manan P. Shah[2], and Irving Biederman[1,2]

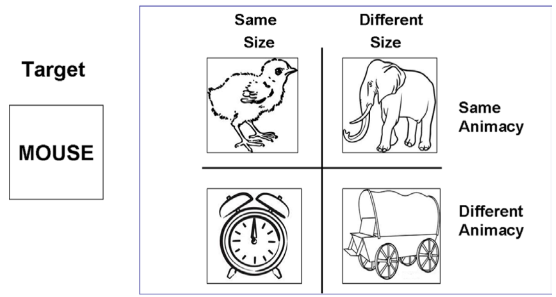
[1]Department of Psychology, [2]Neuroscience Program, University of Southern California (*xiaokunx@usc.edu)

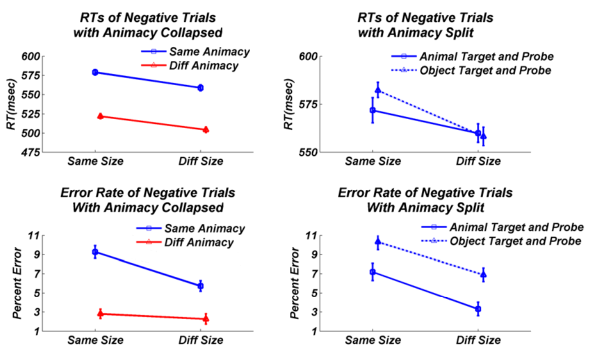
Automatic Coding of Object Size and Animacy

If you are judging whether a picture is that of a mouse, even though the task requires processing neither animacy or size, it is markedly more difficult to respond NO when the picture matches the target in animacy, e.g., elephant, or size, e.g., alarm clock.

Subjects performed a ***verification*** task in which they judged whether a ***picture*** matched a preceding target ***word***. Interest centered on negative trials where the picture could differ from the target in animacy and/or referential size.

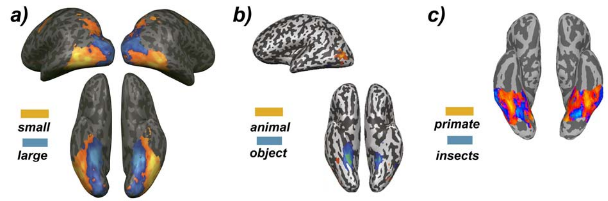
Non-match (negative) trials





Sizable ***costs*** when the negative picture matched the target word in size or animacy! Costs were approximately additive with the costs of size congruency smaller that that for Animals ($p = 0.16$) than that of Objects ($p < 0.001$).

Prior research: Separate studies on Size or Animacy

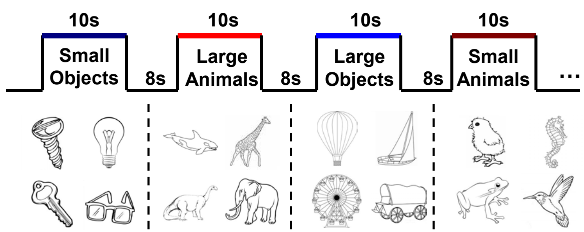


a) Size Map: large vs. small reference size of inanimate objects (Konkle & Oliva, 2010)

b) Animacy Map: animal vs. non-living objects. (Mahon et al., 2009)

c) Multi-voxel gradient of animacy: from insects to birds to primates. (Connolly et al. (2012)

What Map Emerges when Size and Animacy are Varied in Concert?



Blocked presentations of objects (Small-Large Crossed w. Animals-Objects). Subjects performed an oddball (red bounding box aroud image) detection task. Mean referential size of was comparable between animals and objects (Object size from Konkle & Oliva, 2011; Animals fromWikipedia).

Reference:

Connolly, A., Guntupalli, S., Gors, J., Hanke M., Halchenko., O. Y., Wu, Y. C. Abdi., H., & Haxby, J. V. (2012). The Representation of Biological Classes in the Human Brain. *J. Neurosci. 12*:2608-2618.

Mahon BZ, Anzellotti S, Schwarzbach J, Zampini M, Caramazza A (2009) Category-specific organization in the human brain does not require visual experience. *Neuron 63*:397– 405..

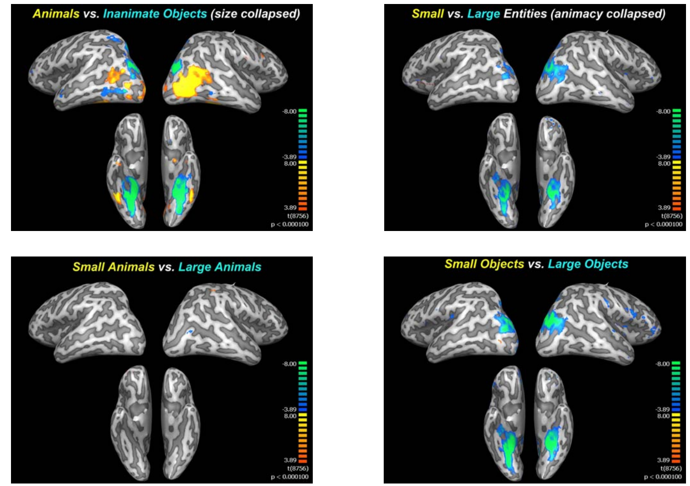
Epstein, R.A., Higgins, J.S., and Thompson-Schill, S.L. (2005). Learning places from views: variation in scene processing as a function of experience and navigational ability. *J. Cogn. Neurosci*:*17*, 73–83.

Caramazza A, Shelton JR (1998) Domain-specific knowledge systems in the brain the animate-inanimate distinction. *J. Cogn. Neurosci 10*:1–34.

Epstein R, Harris A, Stanley D, Kanwisher N (1999) The parahippocampal place area: recognition, navigation, or encoding? *Neuron 23*:115–125.

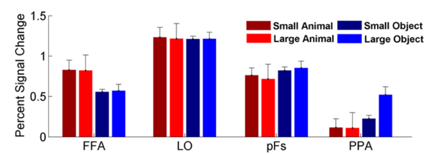
Konkle, T., & Oliva, A. (2012). A Real-World SizeOrganization ofObject Responses in Occipitotemporal Cortex. *Neuron 14*:391–412.

Supported by NSF BCS 04-20794, 05-31177, 06-17699 to I.B.



Animacy: Animals activated ventral lateral temporal cortex, while inanimate objects activated more medial regions of the ventral temporal lobe, and dorsal occipital lobe.

Size: The size map was largely a function of inanimate objects (Konkle & Oliva, 2012), with PPA and transversal occipital sulcus (TOS) preferring large objects. No size map for animals was apparent.



These fMRI effects are compatible with the behavioral task: the effects of animacy were larger than that for size and the size-congruency effect in RTs was significant only for objects.

Why were there stronger size effects, both in behavioral and in fMRI, for inanimate than animate entities?

We tend to interact with large and small objects in fundamentally different ways. Larger objects evoke navigation, producing (perhaps) PPA and TOS activation (Epstein et al., 2005). Small objects (keys, tacks, cups) are manipulable. No significant small object area was found in the present investigation but posterior parietal activation for such stimuli was observed in Konkle & Oliva (2012).

Animate entities, whether large or small, tend to share more features than inanimate objects: they tend to be alive and are capable of self-movement, have eyes / legs / mouth, etc..

Each visual entity defines a network of semantic associations for a number of attributes. For neural efficiency, entities that share many attributes tend to cluster together. Size is just one of many attributes.