



HAL
open science

Investigating semantic properties of objects in scenes using fine-grained crowd-sourced and computational methods

Marek A Pedziwiatr, Sophie Heer, Peter Bex, Antoine Coutrot, Melissa Le-Hoa Võ, Isabelle Mareschal

► To cite this version:

Marek A Pedziwiatr, Sophie Heer, Peter Bex, Antoine Coutrot, Melissa Le-Hoa Võ, et al.. Investigating semantic properties of objects in scenes using fine-grained crowd-sourced and computational methods. 46th European Conference on Visual Perception (ECVP 2024), Aug 2024, Aberdeen, United Kingdom. pp.135. hal-04717006

HAL Id: hal-04717006

<https://hal.science/hal-04717006v1>

Submitted on 1 Oct 2024

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Investigating semantic properties of objects in scenes using fine-grained crowd-sourced and computational methods

Marek A. Pedziwiatr^{1,2}, Sophie Heer¹, Peter Bex³, Antoine Coutrot⁴, Melissa Le-Hoa Võ⁵ & Isabelle Mareschal¹

¹ School of Biological and Behavioural Sciences, Queen Mary University of London, London, United Kingdom

² Centre for Brain Research, Jagiellonian University, Kraków, Poland

³ Psychology Department, Northeastern University, Boston, United States of America

⁴ LIRIS, CNRS, University of Lyon, France

⁵ Department of Psychology, Scene Grammar Lab, Goethe University Frankfurt, Frankfurt am Main, Germany

For most man-made environments, it is straightforward to tell what objects are likely to appear in them and what are not. In the studies investigating the effects of these regularities on human perception, these objects are referred to as, respectively, semantically consistent and inconsistent with a given scene context. While this distinction is useful, its dichotomous nature hampers the investigation of the effects of graded object-context consistency (e.g. a wristwatch and a colander might be consistent with the context of a kitchen table to different degrees). There are several ways to quantify the semantic consistency in a continuous (and not dichotomous) fashion and thereby overcome this limitation. Here, we tested them. In our study, participants viewed scenes with objects that were – according to previous studies – either semantically consistent or inconsistent with these scenes and rated on a six-point scale either the semantic consistency of these objects or the frequency with which they appear in the given contexts. In one experiment, a single group viewed scenes with both consistent and inconsistent objects while in the other, different groups viewed scenes of these two types. We found that although raters in the first experiment tended to select more extreme points from the scale than in the second, the frequency and consistency ratings were always highly correlated (~0.8). Next, we tested if human ratings can be emulated using a pipeline in which GloVe – a computational model capturing semantic relationships between words – processes the names of objects in scenes. The highest correlation between the model consistency scores and the human data amounted to 0.39. We conclude that although the way in which semantic consistency is measured in rating studies affects their outcome, people largely agree regarding how they judge object-scene relationships. Developing computational methods that would emulate these judgements, however, remains challenging.