

## Membership kind and format modality in schema.org categorization

Francesca Zarl, Walter Gerbino

Psychology Unit “Gaetano Kanizsa”, Department of Life Sciences, University of Trieste, Italy

Keywords: categorization, classes, membership, ontology, schema.org.

Ontologies in computer science are conceptual models that are augmented by formal axioms [1]. Schema.org is an example of a web ontology that aims to represent goods online, by defining a shared specification of conceptual elements for data interoperability [2]. Historically, the design of such ontologies has been predominantly governed by membership-based rules. Recently, the consideration of the human cognitive functioning is becoming increasingly significant in ontologies concepts representation [3]. As emerged from the literature analysis, typicality [4-7] and format [8,9] should be considered in ontology concepts representation.

Based on that, the aim of this study is to evaluate participants' responses regarding the concepts representation in schema.org, considering the membership kind (central vs borderline) and format modality (text vs image).

We ran a computer-based experiment in which participants had to categorize random instances (represented using either text or image formats), by answering *Yes* or *No* to an ISA question of the form “*IS exemplar x An instance of category c?*”. After the *Y/N* answer the participant evaluated his/her confidence on a 1-4 scale (with 1= *absolutely certain* and 4= *absolutely uncertain*).

Eleven categories were randomly selected from schema.org. For each category, a balanced set of questions was generated according to a Membership (2) × Format (2) within-subject design. Exemplars might be either central or borderline (the two levels of Membership), and either an image or text (the two levels of Format). The aim was to measure the influence of these variables on categorization.

First, we ran separate oneway analyses of responses to ISA questions using Membership and Format as independent variables. *Y/N* and 1-4 confidence rating data were analyzed according to the unequal variance SDT model. A significant effect of Membership emerged, confirming that ISA judgments were more accurate for central than borderline instances; whereas there was no effect of Format. The response bias was statistically significant, though small, in both analyses: participants showed a liberal bias in the Membership analysis (higher yes rate for borderline than central exemplars) and a smaller bias in the Format analysis (higher yes rate for image than text).

Since the response bias was small, a Membership (2) × Format (2) repeated-measures ANOVA was carried out on percent correct. We obtained a significant interaction: answering ISA questions involving borderline instances were equally difficult for image and text instances (no effect of Format). On the other hand, questions involving central instances was easier for images than text.

1. Gruber, T. R. (1993). A Translation Approach to Portable Ontology Specifications. *Knowledge Acquisition*, 5(2), 199–220.
2. Hepp, M. (2006). Products and services ontologies: A methodology for deriving OWL ontologies from industrial categorization standards. *International Journal on Semantic Web and Information Systems*, 2(1), 72–99.
3. Stark, J., & Esswein, W. (2012). Rules from Cognition for Conceptual Modelling. In P. Atzeni, D. Cheung, & R. Sudha (Eds.), *Conceptual Modeling - ER 2012* (pp. 78–87). Springer Berlin Heidelberg.
4. Au Yeung, C. M., & Leung, H. F. (2006). Ontology with Likeliness and Typicality of Objects in Concepts. In D. W. Embley, A. Olive, & S. Ram (Eds.), *ER 2006* (Vol. 4215 LNCS, pp. 98–111). Springer Berlin Heidelberg.
5. Au Yeung, C. M., & Leung, H. F. (2010). A formal model of ontology for handling fuzzy membership and typicality of instances. *Computer Journal*, 53(3), 316–341.
6. Aimé, X., Fürst, F., Kuntz, P., & Trichet, F. (2010). Improving the efficiency of ontology engineering by introducing prototypicality. *Frontiers in Artificial Intelligence and Applications*, 222, 13–23
7. Frixione, M., & Lieto, A. (2014). Towards an extended model of conceptual representations in formal ontologies: A typicality-based proposal. *Journal of Universal Computer Science*, 20(3), 257–276.
8. Lupyan, G., & Thompson-Schill, S. L. (2012). The evocative power of words: activation of concepts by verbal and nonverbal means. *Journal of Experimental Psychology: General*, 141(1), 170–86.
9. Kannan, A., Talukdar, P. P., Rasiwasia, N., & Ke, Q. (2011). Improving Product Classification Using Images. In *IEEE 11th International Conference on Data Mining (ICDM)*, (pp. 310-319), Conference Publishing Services.