

## Theories of Concepts and Contradiction Acceptance

Francesca Zarl<sup>\*1</sup>, Danilo Fum<sup>2</sup>

<sup>1</sup> University of Trieste, Department of Life Sciences, Psychology Unit *Gaetano Kanizsa*, Trieste, Italy

<sup>2</sup> University of Trieste, Department of Humanistic Studies, Trieste, Italy

\*francesca.zarl@phd.units.it

### Abstract

The paper discusses the Heterogeneity Hypothesis about concepts (Machery, 2009) and the empirical support on which it is based. Two experiments are presented which investigate one of the main predictions of the theory, i.e., the fact that people should be willing to accept apparently contradictory sentences about the same entity.

**Keywords:** Concepts; Categorization; Heterogeneity Hypothesis; Hybrid theory; Contextual Framing.

### Introduction

The debate on the acquisition, representation and use of knowledge about categories and concepts has been a constant hallmark of psychological research for almost a century. The discussion was particularly animated during the '70s giving raise to opposing theories. Concepts were considered as variously represented by rules, prototypes or exemplars (Smith & Medin, 1981). The attempt to define which theory could best explain the empirical findings proved vain, even when data from neuropsychological research and neuroimaging were taken into account. As argued by Murphy (2002), it seems now clear that all these theories are, to a greater or lesser extent, wrong and that new ways of thinking about this issue are required. These could be represented by the so-called hybrid theories of concepts or by more radical positions such as the Heterogeneity Hypothesis (HH), put forward by Machery (2009), which casts doubt on the usefulness of the very notion of concepts within psychology.

In the paper the essential features of the HH and its empirical validation are illustrated. Two experiments are then presented which investigate one of the main ramifications of the HH: the fact that people should be willing to accept apparently contradictory sentences about the same entity. Experiment 1 replicates the procedure utilized in Machery and Seppälä with the goal to establish whether people could hold multiple concepts for the same category. In addition Experiment 2 tests the possibility that this effect could be modulated by the conceptual framework adopted in evaluating a sentence, an idea already put forth by Hampton, Dubois, and Yeh (2006) who were however unable to corroborate it.

### The Heterogeneity Hypothesis

Heterogeneity Hypothesis (Machery, 2009) challenges the dominant conceptions, which assume that all concepts, however represented, share the same set of properties. These

properties are used to explain our high-level cognitive skills: how we categorize, reason inductively, draw analogies, etc. It is known that each cognitive process can be variously realized. In fact, there are different ways to reason inductively, to categorize, or to make analogies. The dominant conceptions (what Machery calls the *Received View*) take for granted that these activities are based on the same type of knowledge. Such being the case, it is however difficult for them to explain where this diversity originates from.

The HH assumes that, for each category, there exist different kinds of concepts that have little in common beyond coreferentiality (i.e., the fact that they refer to the same entity). Empirical evidence suggests the existence, for each category, of at least three separate types of concepts capable of storing knowledge of different kinds: prototypes, sets of exemplars and the so-called *theory-theories*. Prototypes represent typical or cue-valid properties of a concept. Furthermore, a concept can be represented (extensionally) through the set of its exemplars. Finally, the *theory-theories* analyse a concept in terms of a causal, nomological or functional theory about the members forming the concept extension. Thus, according to the HH, different representations for the same category (i.e., different concepts) can simultaneously coexist and be used (in a many-to-many relation) by different cognitive processes giving rise to the variety with which they occur.

So-called hybrid theories provide a different explanation for this diversity. These theories claim that each category is represented by a single concept, which can be composed by different parts. Hybrid theories also assume that the distinct parts of a concept can be employed in different forms of the same cognitive process. For example, Osherson and Smith (1981) claim that a concept is composed by two parts: a set of properties, that are necessary and sufficient to define the concept, and a prototype. These components come into play in two distinct form of categorization: one based on prototype similarity and the other on definitions.

Different predictions about the output of the cognitive processes could be derived by these assumptions. For the hybrid theories, the separate representations work together contributing to a consistent result. On the other hand, for the HH, each concept is involved in a totally separate process and could lead to conclusions that may be uncoordinated or even contradictory.

Machery and Seppälä (2011) empirically investigated these predictions through an experimental procedure, which required the evaluation of pairs of contradictory sentences. Participants were asked to establish (on a seven-

point Likert scale) how much they agreed with affirmative or negative classification statements. According to the HH, if participants accept two mutually contradictory sentences as true, they should be able to utilize two different concepts for the same category. It was further hypothesized that contradiction acceptance would be higher for sentences allowing the use of conflicting membership criteria (the *target* sentences) in comparison with those for which the criteria coincide (the *control* ones). In other words, because tomatoes are technically fruit but share many properties with vegetables, it would be possible to accept both: “*In a sense tomatoes are vegetables*” and “*In a sense tomatoes are not vegetables*”. On the other hand, control sentences like “*In a sense lions are animals*” and “*In a sense lions are not animals*” should not be considered simultaneously true due to the fact that lions are in fact animals and that they are also typical animals. So the HH allows contradiction acceptance, at least for certain kinds of sentences. On the other hand, hybrid theories, which do not allow the existence of multiple representations for the same category, would have difficulty in explaining the phenomenon.

Machery and Seppälä (2011) obtained findings, which are actually compatible with the idea that people could hold multiple concepts for the same category and thus support the HH. In their first experiment, for example, the average percentage of agreement with both sentences was 27.9% for the target pairs and 2.78% for the control ones. The evidence is however not conclusive due to the insufficiency of their statistical analyses and to some limitations in the experimental procedure. In order to further investigate this issue, we carried out the following experiment.

### Experiment 1

The experiment aims to replicate the first study reported in Machery and Seppälä (2011). In comparison with the original experiment, greater attention is given to the internal validity of the theoretical constructs, to the control over the experimental material, and to the evaluation of the results through stricter statistical tests. Basically, the experiment asked participants to determine how much they agreed with the statements contained in pairs of contradictory sentences. The major differences with Machery and Seppälä (2011) consisted in the introduction of a new type of control sentence and in the fact that participants were divided in two groups. To one group (called *Pair*), the contradictory sentences were presented in pairs and participants were asked to evaluate them sequentially. To the other group (*Single*) every sentence was presented, and had to be evaluated, separately. This avoids that the assessment could be affected by the evaluation given to the other pair.

### Method

**Participants:** 40 participants (31 females) from different socio-cultural background took part in the experiment. They were all from the Trieste area and their age varied from 20 to 42 years (mean=26.1, sd=4.8). Participants were

randomly assigned to the two experimental conditions (*Pair* vs. *Single*).

**Materials:** In the experiment, 32 pairs of sentences were used. In each pair, a sentence made an affirmative statement and the other contained its negation. Four different kind of sentences were used:

- T1: The subject of the sentence was similar to the prototype concept of the sentence predicate but it did not belong to its extension (e.g., “*In a sense bats are birds*”);
- T2: The subject was dissimilar to the predicate prototype but, in fact, it was an atypical member of its extension (e.g., “*In a sense penguins are birds*”);
- T3: The sentence subject was both similar to the predicate prototype and it was a member of its extension (e.g., “*In a sense canaries are birds*”).

Sentences T1 and T2 correspond to the *Target* sentences utilized by Machery and Seppälä (2011), while T3 correspond to their *Control* sentences. In the experiment, to fully balance the membership and typicality factors, we utilized another kind of statements:

- T4: The subject was both dissimilar to the predicate prototype and did not belong to its extension (e.g., “*In a sense toads are birds*”).

As it is apparent from the above description, the subjects of sentences T1 and T2, the *Target* sentences which allow the use of different evaluation criteria (typicality vs membership), are borderline members of their natural categories, while in sentences T3 and T4, which constitute the *Control* ones, they are typical members of them. Subjects and predicates of the sentences were balanced within each kind of sentence in order to obtain all their possible combinations.

As a consequence of the criteria adopted in the construction of the materials, in T1 and T4 are true the negative sentences, while in T2 and T3 are true the affirmative ones.

**Design:** We adopted a 2x2 mixed design, having the modality of sentence presentation (*Pair* vs *Single*) as a between subjects factor and the kind of sentence (*Target* vs *Control*) as a factor within.

**Procedure:** All the sentences (64 in total) were gathered in a leaflet whose pages contained eight sentences each. Next to each sentence was printed a line 7 cm long whose extreme points were marked with the labels “*Completely disagree*” and “*Completely agree*”, respectively. Participants had to indicate the degree of their agreement with the sentence by putting a vertical mark on the line. For participants in the *Pair* condition, each page included four randomly chosen pairs of sentences. In each pair, the order of presentation of the positive and negative sentence was randomized. For the participants in the *Single* group, each leaflet page contained eight different sentences drawn randomly from the total of 64 possible ones.

To facilitate the comparison with the data reported in

Machery and Seppälä (2011) participants' responses were translated into a seven-points Likert scale with marks comprised between 0 and 1 cm scored as 1, marks comprised between 1 and 2 cm scored as 2 etc.

## Result

As discussed above, the fundamental difference between the predictions made by the HH and the hybrid theories concerns the degree to which participants are willing to accept mutually contradictory sentences. In fact, it is not easy to establish when this happens. Machery and Seppälä (2011) used two dependent variables: (a) the percentage of participants giving an answer greater than or equal to 4 to both sentences of a pair, and (b) the absolute value of the difference between the answer given to the positive and the negative sentence of each pair. The rationale behind these measures is quite clear. To accept a contradiction it is necessary that both the positive and the negative sentence be considered true. Moreover, the difference between the scores should be small. For instance, if a member of a pair of sentences receives a score of 7 (corresponding to a "Completely agree") and the other member a score of 4 (corresponding, more or less, to "not sure"), we are not licensed to assume that the participant holds both sentences as true.

From our point of view, the criteria used in the previous study are not entirely satisfactory in order to establish when a real contradiction is present. In addition to the difference between the scores given to the sentences of a pair (which we will call *Delta*), the absolute value of the scores should be considered, too. In other words, there is a difference between a *Delta*=1 deriving from the scores of 5 and 4 and a *Delta* of the same magnitude resulting from a 7 and a 6. The former denotes a situation of uncertainty while the latter indicates a real contradiction. In addition to the dependent measures used in the previous study we therefore took into account a new variable: the *Sum* of the scores assigned to the two sentences. Contradictory sentences are, therefore, characterized by a small *Delta* and a high *Sum*, indicating that the participant was pretty confident about their truth.

Due to space limitation, we report here only the main finding obtained from the experiment, i.e., the number of contradictions accepted by participants in the two experimental conditions for the different types of sentences. To obtain this result, in analogy with Machery and Seppälä (2011), we discarded from the analysis the pairs whose true sentence obtained a score less than 4. This allows to take into account the fact that, after all, some people could ignore the natural superordinate class of a given concept and actually believe, for instance, that "In a sense carrots are fruit". Of the remaining sentences, we considered accepted as contradictory those pairs whose *Delta* was less than or equal to 2 and whose *Sum* was higher than 10. Figure 1 reports the average of accepted contradictory pairs for each sentence type in the two experimental conditions.

## Discussion

The findings seem at first sight compatible with the HH. Similarly to what was obtained in Machery and Seppälä (2011), participants actually seem to be willing to accept contradictory statements, with acceptance higher ( $F_{1,38}=5.733$ ;  $p=0.022$ ) for the *Target* sentences (T1 and T2) than for *Control* ones (T3 and T4). However, there are some important reservations to be made.

First of all, it should be observed that the amount of real contradictions in the *Pair* condition is much lower than that expected according to the previous studies. Moreover, the ANOVA revealed a significant difference ( $F_{1,38} = 12.04$ ;  $p = 0.001$ ) between the *Pair* and *Simple* conditions. The fact that people were more willing to judge two contradictory sentences as true when they were presented one at a time, suggests, however, that contradiction acceptance could be better explained by the access to different aspects of the same hybrid representation than by the simultaneous activation of different concepts.

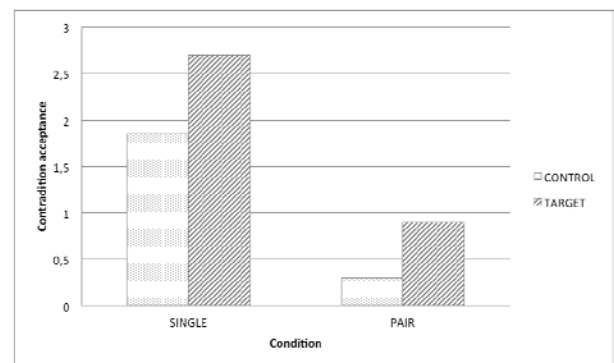


Figure 1: Average sentence contradiction acceptance.

This idea is based on an observation made several times (e.g., Braisby & Franks, 1997) that a major source of variability in categorization judgments derives from the absence of an explicit context for the task. The fact that all our sentences, following Machery and Seppälä (2011), are introduced by "In a sense..." does not help to establish a clear framework for categorization. It is therefore reasonable to believe that, when the contradictory sentences are presented in pairs, they are evaluated within the same framework (hence the lower number of contradictions). On the other hand, when they are presented separately, participants are free to choose every time the perspective through which the sentences are evaluated (giving thus rise to a higher number of contradictions). To assess the reliability of this hypothesis we carried out Experiment 2.

## Experiment 2

Several studies on classification manipulated the context in which the task had to be performed (Murphy, 2002). Closest to our purpose is the work carried out by Hampton et al. (2006). In their experiment, the instructions given to participants contrasted a purely pragmatic classification context with a more technical one, and these were compared

with a no-context control condition. Contrary to the expectations, none of the dependent measures was influenced by the context.

In Experiment 2 we investigated whether contradiction acceptance could be reduced by providing participants with a clear purposive context for sentence evaluation. It could be speculated that such a context will encourage the adoption of a homogeneous membership criterion, thus reducing the need to resort to multiple concepts for the same category.

## Method

**Participants:** 60 University of Trieste students (48 females), whose age varied from 18 to 53 years (mean=23.1, sd=8.0), participated to the experiment. They were randomly assigned to three experimental conditions called *Neutral*, *Prototype*, and *Theory*.

**Materials:** The instructions provided to participants differed in the context according to which they were asked to evaluate the sentences. For the *Theory* group the instructions highlighted the fact that concepts are structured according to a taxonomy based on strict membership rules. Participants in the *Prototype* condition were said that concepts are related according to their similarity. Finally, the instructions for the *Neutral* group did not provide any specific indication about the context to be adopted.

In the experiment the same sentences of Experiment 1 were used. In the case of *Theory* group, however, they were introduced by the expression “*In a technical sense...*” while each sentence of the *Prototype* condition began with the words “*According to common sense...*”. Sentences of the *Neutral*, like those of the *Single* condition of the previous experiment, started with “*In a sense...*”.

**Design:** The experiment followed a 3x2 mixed design. Participants were assigned to three experimental groups (*Neutral*, *Prototype* and *Theory*) which constituted the between subjects factor, while the sentence kind (*Target* vs *Control*) was the factor within.

**Procedure** The procedure was identical to that employed in the *Single* condition of Experiment 1.

## Result and discussion

Figure 2 reports the average of accepted contradictions for the different kinds of sentences in the separate experimental conditions. The first thing to note is that the results in the case of the *Neutral* context are similar to those obtained in the *Single* group of Experiment 1. In fact the conditions in the two experiments were identical and no significant difference was obtained between their results.

In Experiment 2 a mixed ANOVA revealed the main effect of the different type of context ( $F_{2,57} = 3.966$ ;  $p = 0.024$ ) on the number of contradictions accepted by participant while no difference between the *Target* and *Control* sentences was found. A Tuckey HSD post-hoc test

revealed a significant difference between the *Neutral* and the *Theory* conditions only ( $p = 0.011$ ) while the other comparisons did not yield statistically significant results. Providing a purposive context had therefore the effect to reduce the number of contradictions accepted by participants, making the *Target* sentences similar to the *Control* ones. These results supports the hypothesis that the findings obtained by Machery and Seppälä (2011) and in our Experiment 1 could derive not from the simultaneous access to different concepts for the same category but from the fact that, in case of sentences presented individually, participants may adopt each time a different reference context. If this hypothesis is true, the contradictions that were detected may be more apparent than real and the support for the HH could be weaker than previously thought.

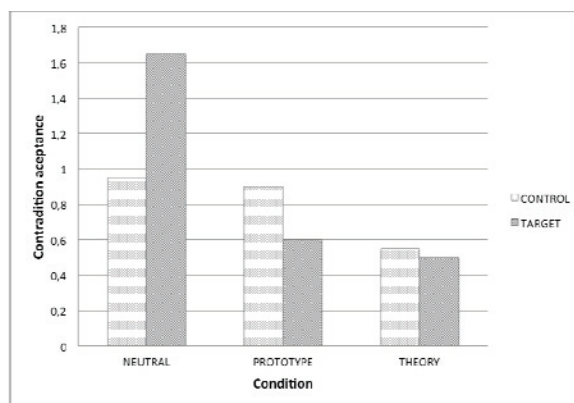


Figure 2: Average sentence contradiction acceptance.

## Conclusion

In this study we reported two experiments, testing the prediction made by the HH and the hybrid theories about the willingness to accept contradictory statements. While the first experiment obtained findings that seemed to corroborate the HH, the second experiment suggests a different interpretation more in line with the assumptions of the hybrid theories. The results of the experiments are however not conclusive because, as it was the case for “classical” theory of concepts, it is difficult to find a critical test that allows to discriminate between them. In any case, it seems clear that the adoption of a purposive context greatly reduces the vagueness and inconsistency in the use of concepts, limiting the need to resort to multiple representations for the same conceptual category. It remains to be determined whether the context could directly influence the process of categorization, an effect that seems dubious in the light of the findings reported in Hampton et al. (2006), or its role should be limited to the evaluation of classification statements.

### Reference

- Braisby, N., & Franks, B. (1997). What does word usage tell us about conceptual content? *Psychology of Language & Communication, 1*, 5–16.
- Hampton, J. A., Dubois, D., & Yeh, W. (2006). Effects of classification context on categorization in natural categories. *Memory & Cognition, 34*, 14319–14436.
- Machery, E. (2009). *Doing without concepts*. New York: Oxford University Press.
- Machery, E., & Seppälä, S. (2011). Against hybrid theories of concepts. *Anthropology and Philosophy, 10*, 97–126.
- Murphy, G. L. (2002). *The big book of concepts*. Cambridge, MA: MIT Press.
- Osherson, D. N., & Smith, E. E. (1981). On the adequacy of prototype theory as a theory of concepts. *Cognition, 9*, 35–58.
- Smith, E. E., & Medin, D. L. (1981). *Categories and concepts*. Cambridge, MA: Harvard University Press.