

# 1 The PACKER Model: Producing Alike or Contrasting Knowledge with Exemplar Representations

This model assumes that people represent categories as exemplars in a multidimensional space, and that generation is constrained by both similarity to members of the category being generated *and* dissimilarity to members of other categories. The model assumes people generate categories that are dissimilar to known categories and have strong within-class similarity.

The similarity between exemplars  $x_i$  and  $x_j$  is computed using Shepard’s law:

$$s(x_i, x_j) = \exp \left\{ -c \left[ \sum_k w_k |x_{ik} - x_{jk}|^r \right]^{1/r} \right\} \quad (1)$$

When prompted to make a generation decision, participants are thought to consider both similarity to examples from other categories as well as similarity to examples in the target category. More formally, the aggregated similarity  $a$  between candidate  $y$  and the model’s stored exemplars  $x$  can be computed as:

$$a(y, x) = \sum_j f(x_j) s(y, x_j) \quad (2)$$

Where  $f(x_j)$  is a function specifying each stored example’s degree of contribution toward generation. Although  $f(x_j)$  may be set arbitrarily, in PACKER it is set according to class assignment. For known members of the target category,  $f(x_j) = \gamma$ . For members of contrast categories,  $f(x_j) = -1 + \gamma$ .  $\gamma$  is thus a free parameter ( $0 \leq \gamma \leq 1$ ) controlling the trade-off between within-class similarity and between-class dissimilarity:  $\gamma = 1$  produces exclusive consideration of same-category members, and  $\gamma = 0$  produces exclusive consideration of opposite-category members. When  $\gamma = 0.5$ , the similarity to contrast categories is effectively subtracted from the similarity to the target category.

The probability that a given item  $y$  will be generated given the model’s memory  $x$  is computed using relative summed similarity values across all generation candidates  $y_i$ :

$$p(y) = \frac{\exp \{ \theta \cdot a(y, x) \}}{\sum_i \exp \{ \theta \cdot a(y_i, x) \}} \quad (3)$$

Where  $\theta$  ( $\geq 0$ ) is a free parameter controlling overall response determinism.