1 Description of the non-variant effect of gamma on generation and assignment probabilities in PACKER

This document refers to the version of PACKER as described in the manuscript submitted to Cognitive Psychology (sometime in 2017; henceforth referred to as the manuscript). In this version, PACKER performs very well in predicting the generation of new category exemplars. In particular, the tradeoff parameter (henceforth γ) actually affects the predictions generated in an expected way. However, when trying to produce predictions of category assignments, varying γ does not affect the predictions in any way. The following section of this document describes why this is occurring.

Equation 2 in the manuscript specifies that the aggregated similarity a between generation candidate y_i and a vector of stored exemplars x can be expressed as:

$$a(y_i, x) = \sum_j f(x_j) s(y_i, x_j) \tag{1}$$

where $f(x_j) = \gamma$ when x_j is a member of a target category and $f(x_j) = \gamma - 1$ when x_j is a member of a contrast category.

We can express Equation 1 here a little differently:

$$a(y_i, x) = \gamma \sum_k s(y_i, x_k) + (\gamma - 1) \sum_l s(y_i, x_l) \tag{2}$$

where k iterates over members of the target category and l iterates over members of the contrast category. To make the algebra a little less tedious we can simplify some of the representations here. Let $t_i = \sum_k s(y_i, x_k)$ and $c_i = \sum_l s(y_i, x_l)$, which allows us to express Equation 2 as:

$$a(y_i, x) = \gamma t_i + (\gamma - 1) c_i. \tag{3}$$

According to the implementation of PACKER as seen in the manuscript, an exponentiated Luce's choice rule is used to find the probability p of generating candidate y_i:

$$p(y_i) = \frac{\exp\{\theta \cdot a(y_i, x)\}}{\sum_i \exp\{\theta \cdot a(y_i, x)\}} \tag{4}$$

expanding $a(y_i, x)$ with Equation 3:

$$p(y_i) = \frac{\exp\{\theta \cdot \gamma t_i + (\gamma - 1) c_i\}}{\sum_i \exp\{\theta \cdot \gamma t_i + (\gamma - 1) c_i\}} \tag{5}$$

Thus far, these equations only apply to predictions of category generation. However, predictions of category assignments require some reconceptualization of these equations, although most of the equations retain the same structure. Specifically, for Equations 1 to 5, y_i now represents the assignment of candidate y to target category i.

Whereas the denominator in Equation 5 is the summation of aggregated similarities over all generation candidates in category generation, category assignment predictions require that the denominator specifies the summation over all category assignments. Specifically, i in the summation term represents each category, as opposed to each generation candidate.

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:

$$s\left(x_i, x_j\right) = \exp\left\{-c\left[\sum_k w_k \left|x_{ik} - x_{jk}\right|^r\right]^{1/r}\right\} \tag{6}$$

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:

$$a(y, x) = \sum_j f(x_j) s(y, x_j) \tag{7}$$

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, $f(x_j) = \gamma - 1$. γ 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:

$$p(y) = \frac{\exp\left\{\theta \cdot a\left(y, x\right)\right\}}{\sum_i \exp\left\{\theta \cdot a\left(y_i, x\right)\right\}} \tag{8}$$

Where θ (≥ 0) is a free parameter controlling overall response determinism.