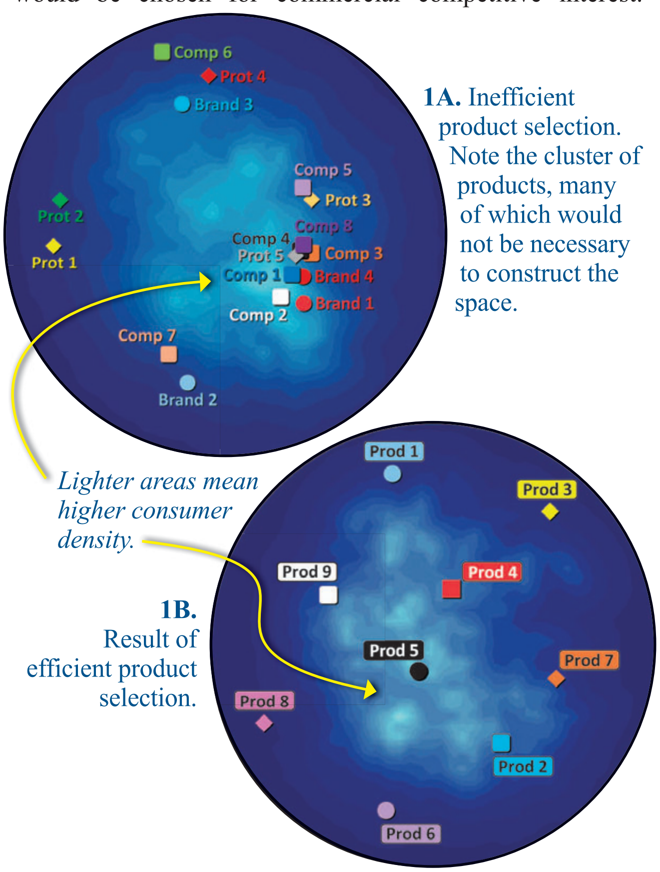
# Selecting Products for a Category Appraisal with Constraints

*Benoît Rousseau, Daniel M. Ennis, and John M. Ennis*

**Background:** Category appraisals are primarily conducted to gain insight into the variables that drive consumer liking but also to understand the relative positions of key products in the drivers of liking space. These appraisals are often quite expensive to field and analyze. Considering the investment in such projects, a critical first step is to carefully consider the selection of the products to include. This selection strongly influences the comprehensiveness, or limitations, of the conclusions reached. When planning a category appraisal, there are often dozens of products to choose from. This occurs because of the number of existing company products, new prototypes and competitor products of interest. Strictly from the standpoint of constructing a reliable Driver of Liking® space, it matters little whether the products chosen for inclusion have high commercial interest to the product developers or consumer insights staff. What matters is whether the products span the space so that the resulting map can be used reliably for computer aided product and brand development. Of course, it would be highly efficient if the products that best span the space are also those that would be chosen for commercial competitive interest.




**1A.** Inefficient product selection. Note the cluster of products, many of which would not be necessary to construct the space.

*Lighter areas mean higher consumer density.*

**1B.** Result of efficient product selection.

**Figure 1A-B.** Landscape Segmentation Analysis® (LSA) maps of products and ideal point densities resulting from selecting inefficient (A) and efficient (B) products.

Figure 1A illustrates the result of a category appraisal in which the primary criterion for product inclusion was the interest expressed in certain company and competitor products by the corporate staff. It can be seen from this figure that there are many similar and clustered products that were tested that would have been unnecessary if the only interest was to develop the drivers of liking space. Figure 1B shows an idealized space well spanned by the products tested. A trade-off now develops between what is necessary to best compute the underlying space and the inclusion of certain products of high interest to the company's staff.

**Scenario:** Your company, in the dairy category, is interested in the current structure of the fruit flavored yogurt consumer landscape. This interest is heightened by the introduction of two new offerings by your main competitors. Your group occasionally conducts category appraisals to better understand competitive threats and to assess your own strengths and vulnerabilities. This information is then used for new product introductions, product and brand positioning and re-positioning. These appraisals generally involve 200 to 300 consumers evaluating sets of 12 to 15 products in a sequential monadic format.

For the current research, budget and timing considerations require a study with a maximum of 12 products (12 samples spread over 3 days of testing per consumer) and 300 respondents representative of the population of interest. Your first task, before initiating the fieldwork, is to choose the most suitable set of 12 products. You have two conditions:

1. The chosen products should span the sensory space to ensure proper representation of the product category characteristics, and
2. Your current offering and the two competitive introductions should be included.

Your task is to select the best group of nine products that, in combination with your current product and the two new competitors, will span the sensory space as well as possible under the constraints. You further accept the limitation that you do not know the sensory drivers of liking in advance and that your choice of products will depend on what is known sensorially about the products under consideration.

One approach is to use sensory profiles from descriptive analysis and visually inspect the first two or three principal components to select products that appear to provide the best spatial spread. For example, from the sensory profiles

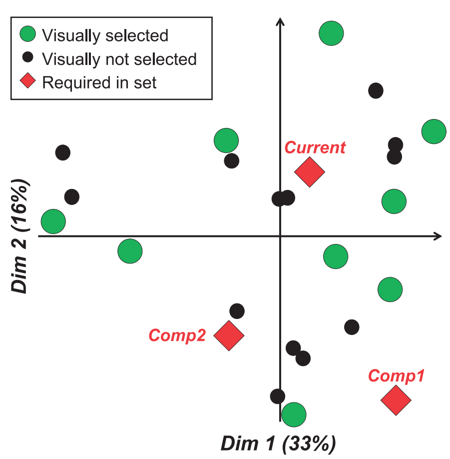



**Figure 2.** First two components of a PCA on the sensory profiles of 25 products.

of 25 fruit flavored yogurts, your first two components are shown in Figure 2. After visual inspection, you would usually choose the nine products color-coded in green on the map. There are also analytic methods for selecting products based on the first two principal components

A limitation of these methods is that they may not provide the most appropriate way to ensure that the sensory space is properly represented. What if there are attributes associated with dimensions different from those you considered for your selection that may be important drivers of liking? Principal components analysis is useful to account for redundancy, not necessarily relevance. Using these methods could result in a suboptimal set of products for a category appraisal. Considering the time and financial investment in this research, there is a need to ensure that the best set of products is selected.

**Graph Theoretic Analysis for Optimal Combinations:** Graph theory, the mathematical study of relationships between items, has seen applications in sensory and consumer science as diverse as finding an optimal mix of pizza toppings, optimizing the compatibility of items in US Army meals ready-to-eat, or selecting product bundles of sparkling fruit juice beverages[1,2,3]. When graph theory is applied within the context of sensory and consumer science, we refer to it as Graph Theoretic Analysis (GTA). In a previous technical report[4] we described how GTA can be used to select optimal sets of products for a category appraisal project without constraints and using solely the first two components of a principal component analysis. Now we extend that idea to the requirement to include pre-selected products which form the constraints as well as to use information from the PCA's multidimensional space.

For this application, to take into account the full dimensionality of the PCA solution, we measure dissimilarity using the product factor scores to calculate the multidimensional Euclidean distances between each pair of products. Table 1 is a subset of the product multivariate distances.

| | P1 | P2 | P3 | P4 | P5 | Comp 2 | |
|---|---|---|---|---|---|---|---|
| **P1** | 0 | 0.40 | 5.09 | 4.13 | 7.50 | 2.40 | … |
| **P2** | 0.40 | 0 | 4.75 | 3.85 | 7.28 | 2.00 | … |
| **P3** | 5.09 | 4.75 | 0 | 1.45 | 3.59 | 3.27 | … |
| **P4** | 4.13 | 3.85 | 1.45 | 0 | 3.54 | 2.95 | … |
| **P5** | 7.50 | 7.28 | 3.59 | 3.54 | 0 | 6.46 | … |
| **Comp 2** | 2.40 | 2.00 | 3.27 | 2.95 | 6.46 | 0 | … |
| **…** | … | … | … | … | … | … | … |

**Table 1.** Pairwise multivariate Euclidean distances between the 25 products over 15 dimensions.

Graph Theory can be used to find smaller collections of related items out of a larger group of many items. When it comes to product selection, we can consider items to be related if they are dissimilar. In this way, the problem of finding a collection of items that are as dissimilar from each other as possible is transformed to the problem of finding a collection of items that are as related to each other as possible – a well-studied problem in graph theory.

**Product Selection for the Yogurt Category Appraisal:** Using GTA applied to the distances from Table 1 you find an optimal set of twelve products that will include your current offering and your two competitors. The analysis actually returns two solutions based on the specifications. Since both solutions are suitable, you pick one of the two based on cost and product availability. Figure 3 shows the nine products chosen in addition to the three required, plotted on the first two components of the PCA. The analysis was not restricted to data on the first two principal components, they are simply used to illustrate the results. As can be seen, the set shows some clear differences from that presented in Figure 2. This choice makes much better use of the information available than the previous analysis based on a limited set of principal components.

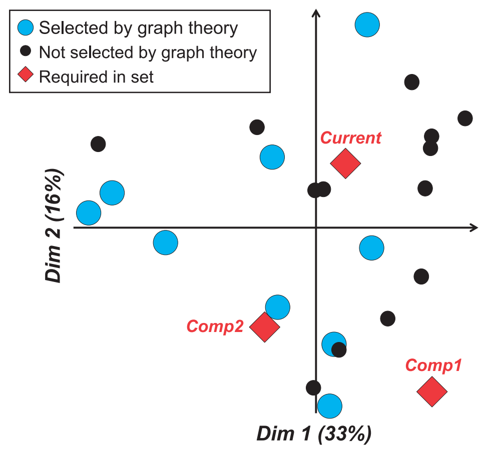



**Figure 3.** Twelve products selected using GTA based on multivariate distances between the products on the complete set of principal components. The first two principal components are shown.

**Conclusion:** The sample selection process is a critical step in planning a category appraisal. If this step is not executed properly, there is a risk of ignoring important underlying sensory differences and missing characteristics relevant to consumer acceptability. Graph theoretical analysis (GTA) permits the selection of a set of products taking into account all of the available sensory information. The method has been adapted to provide the best subset of samples that contains some products that must be included based on required selection criteria.

## References

1. Ennis, J. M., and Corbett, C. (2012). Finding best product-bundles of sparkling fruit-juice beverages using Graph Theoretic Analysis (GTA). In *2012 Sensometrics Conference* in Rennes, France.

2. Nestrud, M., Ennis, J. M., Fayle, C. M., Ennis, D. M., and Lawless, H. T. (2011). Validating a graph theoretic screening approach to food item combinations. *Journal of Sensory Studies,* **26**(5), 331–338.

3. Nestrud, M., Ennis, J. M., and Lawless, H. T. (2012). A group level validation of the supercombinatorality property: Finding high-quality ingredient combinations using pairwise information. *Food Quality and Preference,* **25**(1), 23–28.

4. Ennis, D. M., Rousseau, B., and Ennis, J. M. (2015) *Tools and Applications of Sensory and Consumer Science*. Part 6, Combinatorial Tools, pp. 105-115. Richmond, VA: The Institute for Perception.