How to Diagnose the Need for 3D Unfolding

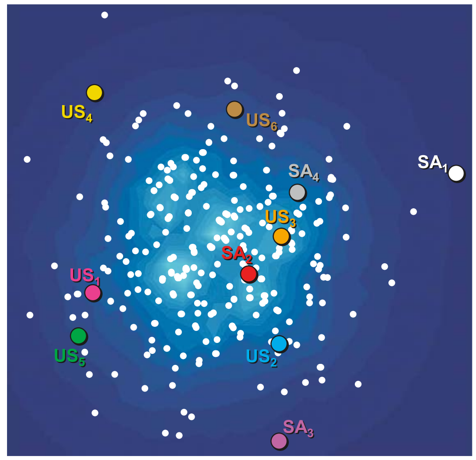
Benoît Rousseau

Background: Multivariate mapping techniques are frequently and commonly used to visualize the large amount of data generated in sensory and consumer testing experiments[1,2]. Since it is desirable to summarize data using as simple a model as possible, multidimensional solutions that capture the relevant information with fewer dimensions are usually prioritized. Moreover, it is less challenging to communicate results in two dimensions. Thus, many analyses are conducted and summarized in two dimensions and this approach is often appropriate. However, using only two dimensions can ignore important and relevant information contained in higher dimensions. In this report, we illustrate how an extra dimension is sometimes needed to capture relevant information when the multidimensional unfolding method, Landscape Segmentation Analysis® (LSA), is applied, so that the proper dimensionality is used to uncover the drivers of liking space.

Scenario: Your company produces several natural orange juice products. While the fruits generally come from Florida, you are investigating other sources from South America, including Brazil, to complement your current sources and potentially reduce costs. There is a need to compare the juices derived from these various oranges to confirm suitability. Your management requests that you recommend the best South American option from a sensory perspective. To do so, you set up a series of analytical investigations, including standard instrumental and internal sensory measurements. In order to capture the consumer's opinion on the products, you also conduct a category appraisal in three locations with a combined number of 300 users who consume orange juice at least weekly. The set of ten products you select contain six products that are representative of your local market as well as four samples from different potential South American suppliers. Once the data becomes available (two days of testing per consumer, five samples per day), you analyze it using LSA. The map that you obtain in two dimensions is shown in Figure 1. The average consumer liking rating for each of the products is also provided.

On the LSA map, products more centrally located within the cloud of ideal points should be liked to a greater extent because they tend to receive the lowest numbers of low liking scores (the distance between an ideal point and a product is inversely related to liking[3,4]).

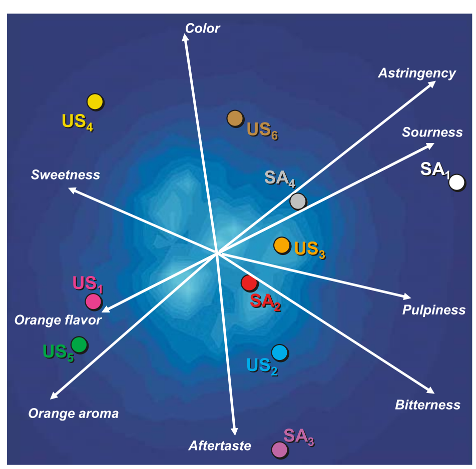
Reviewing Figure 1, you notice something surprising when comparing the average liking ratings to the products' location on the map: While the best liked product (US_3, market leader) is fairly central within the cloud of ideal points, illustrating its higher overall liking rating, two South American products, SA_2 and SA_4, are placed very close to it and could be worth investigating. However, while SA_4 received a high average liking rating, SA_2 did not; in fact, SA_2 was the least liked option. Thus you question the meaningfulness of the solution. Adding sensory descriptive information, as shown in Figure 2, to explain the drivers of liking space does not explain the anomaly. Attributes hypothesized to drive lower likings are not found as drivers of liking. In the presence of apparently conflicting evidence (average liking ratings vs. map locations), you are having difficulty making clear recommendations on which South American product to consider for potential future development.



Product	Mean Liking
US_3	8.08
US_2	7.68
SA_4	7.67
US_1	7.45
US_6	7.26
US_5	7.06
SA_3	6.52
US_4	6.44
SA_1	5.50
SA_2	5.33

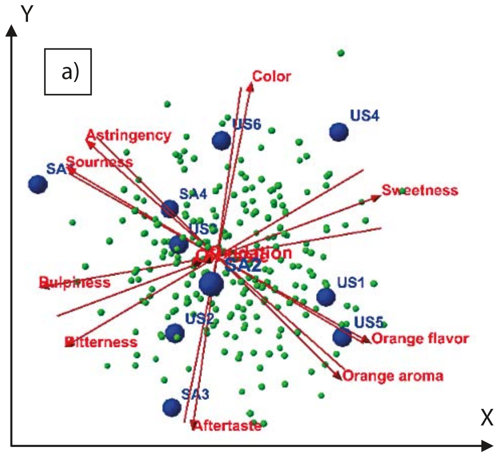
Figure 1. 2D LSA map of the orange juice data showing product positions and consumer individual ideal points with average liking ratings per product on a 9-point hedonic scale. The white dots represent consumer ideals and the products are labeled according to their source (US and SA).

Perceptual Standard Deviations and Dimensionality: LSA can theoretically be conducted in as many dimensions as the amount of data available permits. However, as mentioned previously, limiting the solution to a few dimensions allows a more understandable visual representation of the data. The question is then whether a two-dimensional representation is a suitable summary of the data. Model fitting provides various statistical diagnostics, such as the variance explained in the product means and individual ratings data, but these statistics may not be enough to provide guidance on the best choice of a solution[2].



Scale Label	Correlation
Sweetness	0.99
Pulpiness	0.99
Orange flavor	0.99
Aftertaste	0.99
Color	0.99
Bitterness	0.99
Astringency	0.98
Orange aroma	0.98
Sourness	0.94
Off taste	0.14
Oxidation	0.06

Figure 2. 2D LSA map with drivers of liking. Also shown is a table summarizing the correlation between the product projections onto the sensory directions and the original trained panel sensory information.



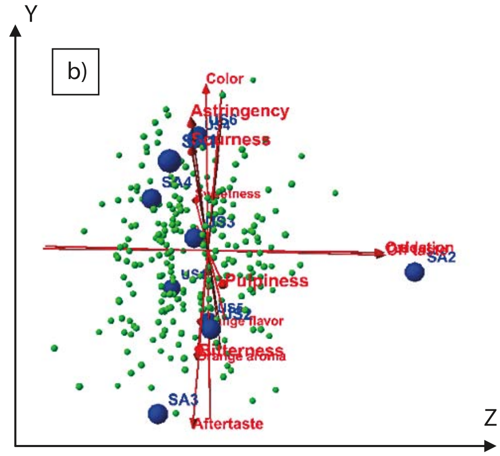


Figure 4. Two views of the 3D LSA output: XY axes (a) and ZY axes (b).

Another piece of information available in conducting an LSA is the products' estimated perceptual standard deviations. Product perceptual standard deviations are linked to the fact that, according to the model, product perceptions vary from moment to moment and across individuals. A relatively large value for a given product will indicate that higher variability is associated with it.

In some instances, a product's estimated perceptual standard deviation can be markedly large compared to those of the other mapped products, and its location may be somewhat counterintuitive based on overall performance (e.g., the product is placed centrally or next to products that perform much better in terms of consumer acceptability). This relatively larger standard deviation could simply be due to relatively greater sensory variability in the product. But it is also possible that this larger deviation indicates that LSA had difficulty fitting the product in two dimensions. Refitting the data by taking into account an additional dimension sometimes results in a lower standard deviation for the product and a more intuitive solution that fits better with other information about the products as well.

3D versus 2D: Thinking that the SA_2 placement might be driven by a third relevant dimension, you look into the products' estimated perceptual standard deviation. You indeed find that this product's estimated standard deviation is much larger than those of the other products (Figure 3).

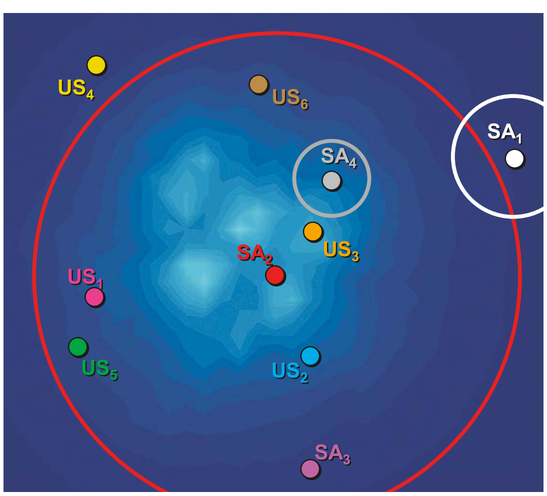


Figure 3. 2D LSA map with products' perceptual standard deviations.

You then re-run LSA in three dimensions and confirm your intuition[5] (Figure 4). While the product's R^2 value stays about the same and very high (~1), the subject R^2 increases from 0.91 to 0.97. More important, you find that the SA_2 product exists in a third dimension explained by the "Off taste" and "Oxidation" attributes, while the other test products are located mainly in a single plane represented in Figures 1 and 2. In fact, the space shown in Figure 4a projects in two dimensions onto a plane very similar to that shown in Figures 1 and 2. You conclude that the original 2D solution was not an accurate representation of the consumer data and that the 3D solution offers a more meaningful interpretation. According to this interpretation, SA_2's perceptual standard deviation is more similar to those of the other products, SA_2 is positioned away from the main cloud of ideal points, and the high scores for SA_2 on the two additional drivers of liking variables - "Off taste" and "Oxidation" - underscore its inherent weaknesses. Therefore, you conclude that SA_2 is not a suitable option and feel confident recommending SA_4 for future development.

Conclusion: With Landscape Segmentation Analysis®, as with any multivariate mapping technique, the question arises as to how many dimensions should be used to best describe the data, without rendering the visual solution difficult to comprehend. While standard model fitting statistics, such as the variance explained by the model, provide insights on the quality of the solution, they may not be enough to lead to optimal interpretation. The products' perceptual standard deviations can also be useful and sometimes suggest the need for additional mapping dimensions. Analyses such as those described in this report may provide valuable information on the best way to analyze and summarize data on consumer hedonics.

References and Notes

1. Meullenet, J., Xiong, R. and Findlay, C. (2007). *Multivariate and Probabilistic Analyses of Sensory Science Problems.* New York, NY: Wiley-Blackwell.
2. Ennis, D. M. and Ennis, J. M. *(In press).* Mapping hedonic data: A process perspective. *Journal of Sensory Studies.*
3. Ennis, D. (2001). Drivers of liking for multiple segments. *IFPress,* **4**(1), 2–3.
4. Ennis, D. and Anderson, J. (2003). Identifying latent segments. *IFPress,* **6**(1), 2–3.
5. All analyses were conducted using *IFPrograms*™ Professional.