The Institute for Perception

For technical reports from The Institute for Perception or for information about short courses, please visit www.ifpress.com or email us at mail@ifpress.com

Reprinted from IFPress (2015) 18(4) 3,4

Predicting Preference from Liking

Benoît Rousseau and Daniel M. Ennis

Background: Average liking ratings of test products or paired preference proportions are often used to guide product development in consumer product companies. When liking ratings are used, the performance of the products are often tested using an analysis of variance and mean comparisons to select one or more products for further consideration. While a statistically significant difference may provide insight on product superiority, quantification of the effect size itself is also often of interest. Paired preference results are particularly intuitive to quantify such an effect size. Some companies use a preference action standard that corresponds to a meaningful measure of superiority, for instance a 55/45 or 60/40 preference split. Preference tests are valuable to compare test results to these thresholds.

A greater number of variables come into play when considering a liking threshold to set an action standard. For instance, the exact structure of a rating instrument, such as a 9-point hedonic scale or a 7-point numerical liking scale, will produce different measures of hedonic difference. A difference of 0.5 on a 9-point word category scale will be different from the same difference on a 7-point numerical scale and may even be different from a 9-point end-anchored scale. Paired preferences are not subject to these types of effects. However, it is not always possible or cost-effective to use paired preferences. In these situations, a sequential monadic presentation may be used and average liking ratings calculated. Converting these ratings into expected preference proportions would provide effect size information that can be referred to a preference action standard. This report will provide an approach to making that conversion based on Thurstonian models of different types of hedonic data.

Scenario: You work for a global beverage product organization with responsibility for energy drinks available in two main markets: the US and Brazil. Due to specific development paths and differences in local regulations on usable ingredients, your main brand's formulation differs between the markets resulting in a high number of ingredients and suppliers that are market-specific. In an effort to standardize your products, your management directs you to investigate the possibility of a single formulation in the two markets. The new formulation should at least be on par with your current brand and main competitor (Competitor A) and be preferred to another smaller player present in both markets (Competitor B). Your team develops two prototypes that can be produced and sold locally (Prototype 1 and Prototype 2). The investigation will therefore involve a total of five products. Due to the high number of possible pairs, you forego a multiple preference test approach and ask your colleagues in both markets to use a sequential monadic design and liking ratings.

Three hundred consumers are used in each location and the results are shown in Table 1. The liking means are provided along with an indication of whether there was a significant difference between product pairs after adjusting for multiple comparisons.

	Product Pair Differences	
Prototype 1	**USA** 1-9 Point Scale	**Brazil** 0-10 Point Scale
Current	**− 0.05** (7.17, 7.22)	**− 0.78** (7.31, 8.09)
Competitor A	**− 0.19** (7.17, 7.36)	**− 0.58** (7.31, 7.89)
Competitor B	**+ 0.80** (7.17, 6.37)	**+ 0.43** (7.31, 6.88)
Prototype 2	**USA** 1-9 Point Scale	**Brazil** 0-10 Point Scale
Current	**+ 0.29** (7.51, 7.22)	**− 0.10** (7.99, 8.09)
Competitor A	**+ 0.15** (7.51, 7.36)	**− 0.10** (7.99, 7.89)
Competitor B	**+ 1.14** (7.51, 6.37)	**+ 1.11** (7.99, 6.88)

Table 1. Mean differences from prototypes (with rating means in parentheses). Significant differences are shown in red.

Based on these results, you recommend Prototype 2 since it fulfills the target requirements in both countries: Significantly better than Competitor B and not significantly different from Competitor A and your current product. Prototype 1, on the other hand, is found to be significantly inferior to Competitor A and to the current formulation in Brazil and is thus not a viable alternative.

Action Standard: In your project, while significant differences were found, how do they relate to the action standard? Your company has historically considered a 60/40 preference split as the action standard for a meaningful preference. You would like to quantify the results in terms of preference splits so that a final decision can be made whether to choose Prototype 2 or to conduct further research with new alternatives. You only have liking ratings at your disposal and the US and Brazil investigations involved different types of scales.

Linking Liking and Preferential Choice: Thurstonian models for scaling sensory intensities are broadly available for difference and rating methods. When using ratings, the model takes into account the differential use of scale categories by identifying the psychological locations of the scale boundaries and estimating the size of the sensory difference in terms of d' values[1,2]. The same approach can be used with hedonic rating data by assuming the existence of an hedonic continuum like one would assume a sensory continuum. This approach provides a way to predict preference from liking data. The process is illustrated in Figure 1. The liking scores for two or more products are first

tallied and used with the appropriate Thurstonian model for rating data to estimate all *d′* values. The next step is to use the idea that a preferential choice can be modelled as an hedonic 2-Alternative Forced Choice decision between two products. Consequently, the *d′* value between two products obtained from the rating data can be linked to the choice proportion in a hypothetical paired preference test. A *d′* of 0.54 estimated using the liking ratings corresponds to a proportion correct of 65% and thus a 65/35 preference split.

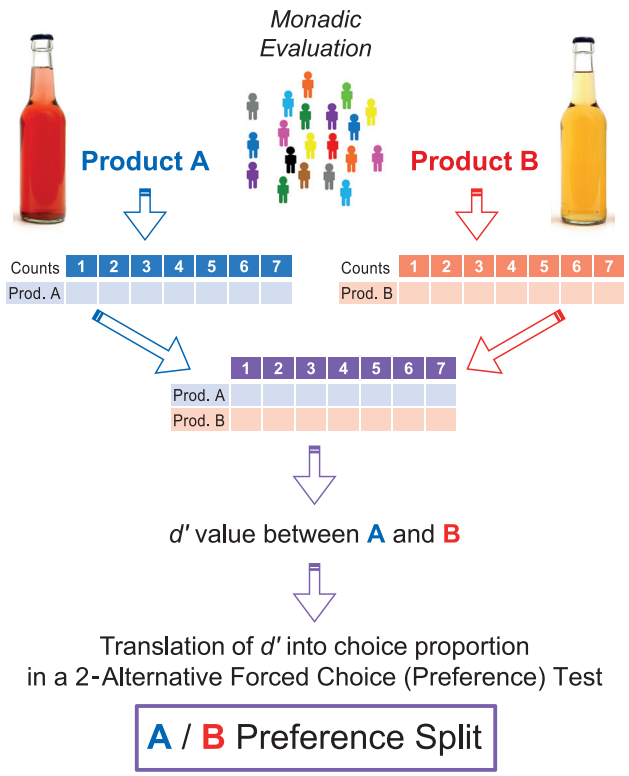

Figure 1. Analytical process to predict preference from liking ratings using Thurstonian modeling.

The value of this approach is that it permits the conversion of liking ratings data into more easily quantifiable preference splits. This theory can be used irrespective of the type of hedonic category rating instruments that are used to generate the liking data (numerical, word, pictorial, or others). An alternative to the idea of an hedonic continuum is to assume that each consumer has an ideal point. Then liking and preference data are modelled and related based on this concept. This idea is the basis for a method called unfolding, of which Landscape Segmentation Analysis (LSA) is an example[3,4,5]. This method is particularly valuable when there is evidence for segmentation.

Predicting Preferential Choice: You reanalyze your data to calculate the underlying *d′* values between all relevant pairs of products and use them to deduce the corresponding predicted preference splits. The results are summarized in Table 2.

Prototype 1	*d′* and Preference Splits	
	USA	Brazil
Current	− 0.02 (49/51)	− 0.31 (41/59)
Competitor A	− 0.09 (47/53)	− 0.24 (43/57)
Competitor B	+ 0.40 (61/39)	+ 0.18 (55/45)
Prototype 2	**USA**	**Brazil**
Current	+ 0.14 (54/46)	− 0.04 (49/51)
Competitor A	+ 0.07 (52/48)	+ 0.03 (51/49)
Competitor B	+ 0.56 (65/35)	+ 0.45 (62/38)

Table 2. *d′* values and predicted preference splits for each product pair (splits above 60/40 shown in purple).

With this new information, you can now confirm that Prototype 2 is indeed a suitable candidate to replace the current formulation. The superiority of Prototype 2 to Competitor B corresponds to a preference split greater than 60/40, while the non-significant differences with the current formulation and Competitor A correspond to preference splits below the action standard. As for Prototype 1 in Brazil, even though it had a significantly greater liking rating than that of Competitor B (Table 1), the predicted preference split (55/45) is lower than the company's historic action standard.

Conclusion: While liking ratings can be used to determine whether products differ in terms of liking to consumers, significant differences are not always enough to reach a decision on whether to go ahead with a product modification. If a preference action standard exists, the use of Thurstonian modeling provides a method to predict preference splits from ratings collected using a monadic or sequential monadic approach.

References (available at www.ifpress.com):

1. Kim, K., Ennis, D. M., and O'Mahony, M. (1998). A new approach to category scales of intensity II: Use of *d′* values. *Journal of Sensory Studies,* **13**(3), 251-267.
2. Ennis, D. M. and Rousseau, B. (2015). A Thurstonian model for the degree of difference protocol. *Food Quality and Preference,* **41**, 159-162.
3. Ennis, D.M., and Rousseau, B. (2004). Motivations for product consumption: Application of a probabilistic model to adolescent smoking. *Journal of Sensory Studies,* **19**(2), 107-117.
4. Rousseau, B., Ennis, D. M., and Rossi, F. (2012). Internal preference mapping and the issue of satiety. *Food Quality and Preference,* **24**(1), 67-74.
5. Ennis, D. M. and Ennis, J. M. (2013). Mapping hedonic data: A process perspective. *Journal of Sensory Studies,* **28**(4), 324-334.