

APPLICATION OF PENALTY ANALYSIS TO INTERPRET JAR DATA – A CASE STUDY ON ORANGE JUICES

Didia Iserliyska^{1,*}, Mina Dzhivoderova², Kremena Nikovska³

¹ Sensory analysis laboratory, Food Research and Development Institute, Vasil Aprilov blvd. 154, Plovdiv – 4000, Bulgaria;

² Department of Technology of tobacco, sugar, vegetable and essential oils, Technological Faculty, University of Food Technologies, Maritza blvd. 26, Plovdiv - 4000, Bulgaria;

³ Department of Nutrition and Tourism, Faculty of Economics, University of Food Technologies, Maritza blvd. 26, Plovdiv - 4000, Bulgaria

Abstract

Penalty analysis is a graphical technique to reveal the possible penalty paid by the product in terms of reduced overall liking by not being Just About Right (JAR) on a characteristic. Thus consumer affective tests were conducted to investigate the use of penalty analysis to model consumer acceptance of six well-known brands of orange juice using the proposed method to infer the drivers of liking from JAR data. Just-about-right (JAR) and hedonic ratings were used to measure each attribute evaluated. Consumers ($n=81$) were asked to rate the overall acceptance using a 9-point hedonic scale. Just About Right (JAR) scales were used to evaluate the rest of the attributes as followed: color, sweet taste, sour taste, bitter taste and amount of pulp. Means and frequencies of each sensory attribute were obtained. Spearman's rank correlation coefficients gave the relationship between the sensory attributes and the overall liking.

Keywords: consumer affective test, hedonic scale, orange juice, penalty analysis, just–about-right (JAR) scale.

1. INTRODUCTION

There are many variations of scales that have used the idea of a central optimal point for the intensity of an attribute. Some authors (Vickers, 1988) used a line marking scale labeled "not nearly sweet enough" at the left end, "just right" at the center, and "much too sweet" at the right end to study the optimization of sweetness in lemonade. Others gave examples of several just right scales to optimize the most important attributes of a product (Pokorny and Davidek, 1986). JAR scales have been used to optimize breads (Bagdi et al., 2016), raisin jams (Rababah et al., 2012), probiotic Petit Suisse Cheese (Esmerino et al., 2013), cooked steaks (Chan et al., 2013), and kefir (Gere et al., 2014), just to name a few.

Penalty analysis is a graphical technique to reveal the possible penalty paid by the product in terms of reduced overall liking by not being "just about right" on a characteristic (Xiong and Meullenet, 2006). The bipolar Just-About-Right (JAR) scales cannot be evaluated using linear approaches for the consumers' ratings are not normally distributed and furthermore the scale has two directions. The steps needed to perform this analysis can be summarized in three points. In case of a five point JAR scale (Gere et al., 2015):

1. Firstly, the JAR values are amalgamated into three groups. Categories 1 and 2, category 3, and category 4 and 5 give the three new levels: “not enough”, “JAR”, and “too much”;
2. The mean overall liking (rating) is calculated for each group. The penalties (or mean drops) are calculated as the differences between the means of the two non-JAR categories and the mean of the JAR category.
3. These values are plotted versus the percentage giving each response in a so called mean drop plot.

The objective of this paper was to determine the acceptance of commercially available orange juice products (a case study) by Hungarian consumers via penalty analysis. This research was conducted to provide directional information for product developers on product reformulation or optimization of the various orange juice products.

2. MATERIALS AND METHODS

2.1. Materials

In our study, six commercially available orange juice samples from the Hungarian market were evaluated by a consumer sensory panel.

2.2. Sample preparation

Sample preparation (stored at 10 °C) was conducted using the same standardized parameters (refrigerator, sample quantity, etc.). Packaging was separated from samples in order to avoid the effect of brand knowledge. The recommendations of Kilcast (2010) were followed during the sample presentation, so the quantities of samples (180 cm³/person in plastic glasses) were prepared by one person using a measuring cup to achieve better homogeneity. Samples were labeled, according to the international practice using 3-digit random numbers and a balanced block design was applied. The samples were presented to the assessors in plastic glasses (200 cm³) at a typical consumption temperature (15 °C), which was strictly monitored to maintain commensurable conditions. Between the evaluations, assessors used a very neutral non-carbonated mineral water as taste neutralizer. Evaluations were performed under artificial daylight-type illumination, temperature control (between 22 and 24 °C) and air circulation.

2.3. Consumer test

One hundred consumers were recruited from the Corvinus University of Budapest, Hungary. Nineteen consumers were left out from the data analysis due to incomplete questionnaires. As a result, the data of 81 consumers were used in the data analysis. Consumers were selected according to relevant market figures: 60%/40% females/males aged between 18 and 30 years, regular orange juice consumers, as they consumed orange juices more than once a week. Consumers were instructed prior to the evaluation to ensure the reliability of the results and asked to evaluate overall liking on a 9-point hedonic scale (1 = “dislike extremely”, 9 = “like extremely”). The attributes of color, sweet taste, sour taste, bitter taste and amount of pulp were evaluated using a nine point “Just About Right” scale.

2.4. Data analysis

In order to conduct penalty analysis, the consumers' overall liking ratings and ratings on the JAR attributes are required. The penalties (or mean drops) are plotted versus the percentage of the consumers giving each response in a so called mean drop plot. Attributes with a large percentage of consumers and large penalties can be found in the upper right quadrant of a plot, providing a quick summary of the most critical diagnostic problems for a product. Penalty analysis was done using XL-Stat Sensory solution (Addinsoft, 28 West 27th Street, Suite 503, New York, NY 10001, USA).

3. RESULTS AND DISCUSSION

3.1. Interpreting the results from penalty analysis

The descriptive statistics for the liking data and JAR variables are shown in Table 1. The correlation matrix displays whether the JAR variables have either “low” or “high” impact on the overall liking and which direction it would be manifested (“too much” or “too little”).

Table 1. Impact of the JAR variables on the overall liking (Spearman's correlation coefficient - product F)

| Variables | Overall liking |
|----------------|----------------|
| Overall liking | 1 |
| Color | 0.044 |
| Sweet taste | 0.264 |
| Sour taste | 0.089 |
| Bitter taste | -0.453 |
| Amount of pulp | 0.040 |

Values in bold are significantly different from 0 with a significance level $\alpha=0.05$

The JAR – variables had low impact on the overall liking ($\alpha=0,05$). The correlation for sweet taste was positive, e.g. the “too much” cases had a lower impact than the “too little” cases and *vice versa* for the bitter taste where the correlation coefficient was negative (Table 1).

For products A and B, JAR – variable for amount of pulp showed low impact on the liking ($\alpha=0.05$) and the correlation turned to be positive (Spearman's correlation coefficient = 0.220 and 0.294, respectively), e.g. the “too much” cases had a lower impact than the “too little” cases (data not presented).

In the case of product C and D, all the JAR – variables have higher impact on the overall liking ($\alpha=0.05$). None of them were significantly different from 0 (data not presented). For product E, JAR – variables for sweet, sour taste and amount of pulp have low impact on the liking ($\alpha=0.05$). The correlation for sweet taste and amount of pulp is positive, e.g. the “too much” cases have a lower impact than the “too little” cases and *vice versa* for the amount of pulp where the correlation coefficient is negative (data not presented).

The following chart visualizes how JAR scores were distributed for each sensory attribute (Figure 1a) and how they merged to a three levels scale (Figure 1b) (Product F).

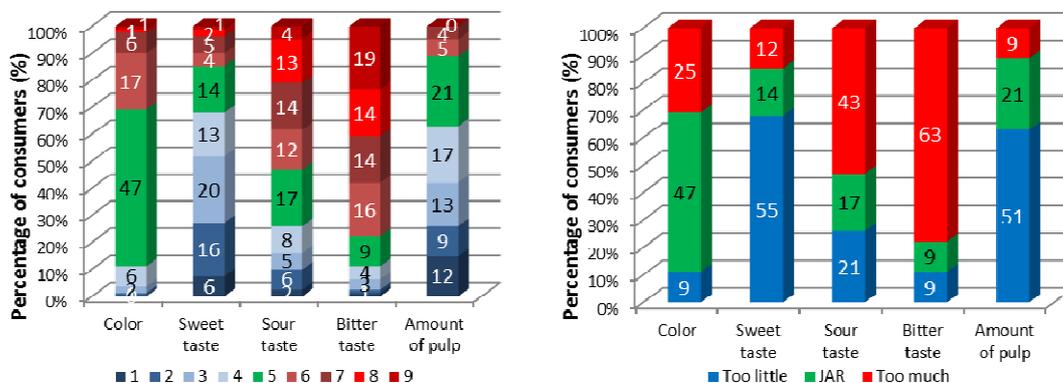


Figure 1. Percentage of panelists (n=81) giving consumer ratings for selected attributes of an orange juice sample (sample F) a) based on 9-point JAR scale; b) based on the collapsed JAR levels

Table 2 corresponds to penalty analysis. The mean drops were calculated for the "too much" and "too little" levels (this is the difference between the overall likings mean for the JAR levels minus the "too much" or "too little" levels. This information is interesting as it shows how many points of liking were lost for having a product "too much" or "too little" for a consumer. The penalty is a weighted difference between the means (mean of liking for JAR - mean of liking for the two other levels taken together).

Table 2. Penalty analysis table (product F)

| Variable | Level | % | Sum(overall liking) | Mean(overall liking) | Mean drops | Penalties | p-value |
|----------------|------------|--------|---------------------|----------------------|------------|-----------|---------|
| Color | not enough | 11.11% | 28.0 | 3.11 | 0.40 | 0.18 | 0.749 |
| | JAR | 58.02% | 165.0 | 3.51 | | | |
| | too much | 30.86% | 85.0 | 3.40 | 0.11 | | |
| Sweet taste | not enough | 67.90% | 163.0 | 2.96 | 2.60 | 2.58 | 0.000 |
| | JAR | 17.28% | 78.0 | 5.57 | | | |
| | too much | 14.81% | 37.0 | 3.08 | 2.48 | | |
| Sour taste | not enough | 25.93% | 60.0 | 2.85 | 2.02 | 1.83 | 0.008 |
| | JAR | 20.99% | 83.0 | 4.88 | | | |
| | too much | 53.09% | 135.0 | 3.14 | 1.74 | | |
| Bitter taste | not enough | 11.11% | 31.0 | 3.44 | 2.66 | 3.01 | 0.001 |
| | JAR | 11.11% | 55.0 | 6.11 | | | |
| | too much | 77.78% | 192.0 | 3.04 | 3.06 | | |
| Amount of pulp | not enough | 62.96% | 164.0 | 3.21 | 0.92 | 0.96 | 0.143 |
| | JAR | 25.93% | 87.0 | 4.14 | | | |
| | too much | 11.11% | 27.0 | 3.00 | 1.14 | | |

For the color and amount of pulp dimensions it is notable that the test is not significant. For the sweet taste dimension the consumers penalize the product when they consider it "not sweet enough". Both mean drops are significantly different from 0, and so is the overall penalty.

Sour taste is strongly penalized by consumers in the "not sour enough" direction. For the bitter taste dimension consumers penalize the product towards "too much bitter". The mean drops are significantly different from 0, and so are the overall penalties.

Figure 2 represents the mean drops plotted versus the percentages of panelists giving responses on each sensory attribute of all the six evaluated products. The plot is divided into four subplots using a vertical line representing 20% of the consumers. The upper right subspace contains the important (more than 20% of consumers' ratings) attributes which have to be emphasized during the product development. In the case of product A the largest percentage of consumers (about 80%) stated out that the amount of pulp was "too little" as well as more than half of the panelists considered the sweet taste as "too little". About 40% felt the color as "too much" and minority (though higher than the 20% threshold set earlier) expressed the bitter and sour taste as "too much" as well. Low percentage of panelists rated the color as "too little" and the sweet taste as "too much", and only few felt the amount of pulp "too much".

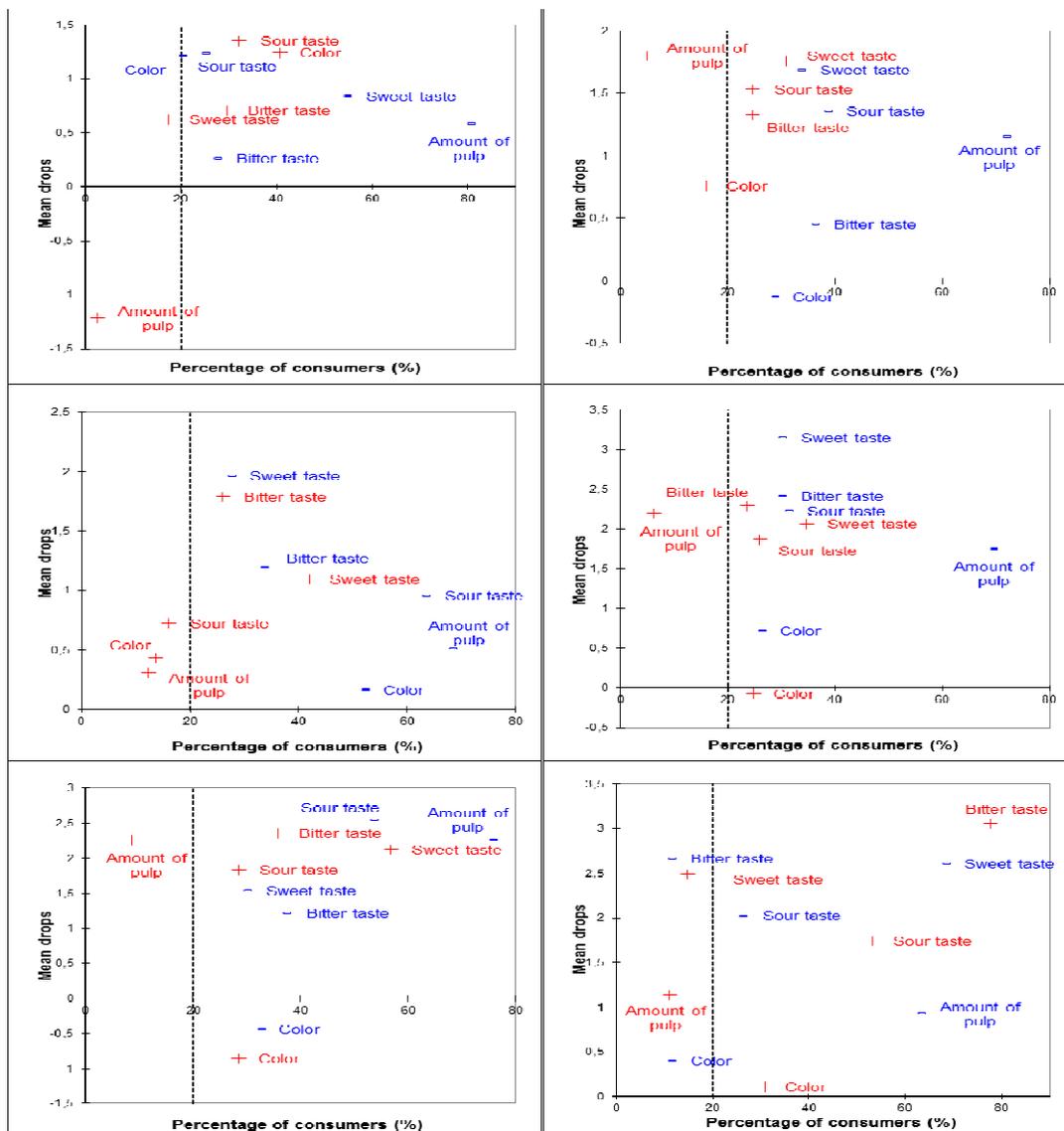


Figure 2. Mean drop plot of all the six evaluated products. Products from top left to bottom right: A, B, C, D, E and F. The too low endpoint of the JAR scales are highlights with red, the too strong endpoint is highlighted with blue. The dashed line represents the 20 % of the consumers.

The amount of pulp was considered by the most of the panelists as “too little”. Less than 40% rate the bitter taste as “too little” and by less than 30% the sour and sweet tastes were “too much” (product B). Smaller percentage rated the bitter taste and color as “too much” and the amount of pulp was “too much” by few. The amount of pulp and the sour taste were considered as “too little” by nearly 70% of the consumers and the color was thought so by 50% (Product C). The sweet taste was rated as “too much” by 40% and about 30% felt like the bitter taste “too little”. Minority (more than 20% though) rated bitter taste as “too much” and the sweet taste “too little”. Those who perceived the sour taste, the color and the amount of pulp as “too much” were negligible. The larger percentage of consumers felt like the amount of pulp “too little”. Bitter and sour tastes were rated by less than 30% as “too little”. Although another 30% considered the color and the sweet taste either “too much” or “too little”. The amount of pulp was “too much” by few (Product D). Majority of the consumers perceived the amount of pulp as “too little”. Larger percentage considered the

sweet taste as “too much” and the sour taste as “too little”. Less than 40% felt like the bitterness either “too little” or “too much” while about 30% think that the sweetness is “too little” and the sourness “too much”. Also the color was thought as “too little” and “too much” by 30%, and the amount of pulp was “too much” by the minority of consumers (Product E). The largest percentage of consumers considers the bitter taste “too much” but sweet taste and amount of pulp as “too little”. About 50% of consumers state the sample is too much sour and small percentage (higher than the 20% threshold set earlier) claim the product is not sour enough. 30% perceived the color as “too much” and less than 20% think that the color and bitter taste are “too little” whereas amount of pulp and sweet taste were “too much” by few (Product F).

4. CONCLUSION

The application of penalty analysis showed that the sensory attribute „amount of pulp” was strongly penalized in the „too little” direction for all the six brands of the selected orange juices by the largest percentage of the consumers. The sweet taste was considered as „too little” for products A&F and „too much” for product E by the majority of panelists. The sour taste was felt like “too little” by nearly 70% of the consumers (product C&E) whereas 50% stated out the color was “too little” (product C) and the sourness “too much” in the case of the product F. The highest mean drop values were computed for the two least preferred products, E&F. Products A, B&C have lower mean drop values which means that those consumers who rated the attributes not JAR, did not penalize the products highly. Penalty analysis showed what are the main reasons of the rejections of products F&E, furthermore, possible improvements of the preferred products were highlighted (increasing the amount of pulp).

5. ACKNOWLEDGEMENT

This research was supported by Erasmus + Programme Key Action 1 Mobility for learners and staff Higher Education Student and Staff Mobility 2014 – 2015 between Sensory Lab in Corvinus University of Budapest, Hungary and Sensory Lab in Food Research and Development Institute, Plovdiv, Bulgaria.

6. REFERENCES

- Bagdi, A., Tóth, B., Lőrincz, R., Szendi, S., Gere, A., Kókai, Z., Tömösközi, S. (2016). Effect of aleurone-rich flour on composition, baking, textural, and sensory properties of bread. *Lwt-Food Science and Technology*, 65, 762–769.
- Chan, S. H., Moss, B. W., Farmer, L. J., Gordon, A., Cuskelly, G. J. (2013): Comparison of consumer perception and acceptability for steaks cooked to different endpoints. Validation of photographic approach. *Food Chemistry*, 136(3–4), 1597–1602.
- Epler, S., Chambers, E., Kemp, K. E. (1998). Hedonic scales are a better predictor than just-about-right scales of optimal sweetness in lemonade. *Journal of Sensory Studies*, 13(2), 191–197.
- Esmerino, E. A., Cruz, A. G., Pereira, E. P. R., Rodrigues, J. B., Faria, J. A. F., Bolini, H. M. A. (2013). The influence of sweeteners in probiotic Petit Suisse cheese in concentrations equivalent to that of sucrose. *Journal of Dairy Science*, 96(9), 5512–5521.
- Gere, A., Losó, V., Györey, A., Kovács, S., Huzsvai, L., Nábrádi, A., Sipos, L. (2014). Applying parallel factor analysis and Tucker-3 methods on sensory and instrumental data to establish preference maps: case study on sweet corn varieties. *Journal of the Science of Food and Agriculture*, 94(15), 3213–3225.
- Gere, A., Sipos, L., Héberger, K. (2015). Generalized Pairwise Correlation and method comparison: Impact assessment for JAR attributes on overall liking. *Food Quality and Preference*, 43, 88–96.
- Kilcast, D. (2010). Sensory quality control for taint prevention. In D. Kilcast, eds, *Sensory analysis for food and beverage quality control* (Part 3 , pp. 156–185). Woodhead Publishing, Cambridge.
- Kovács, Z., Dalmadi, I., Lukács, L., Sipos, L., Szántai-Kóhegyi, K., Kókai, Z., Fekete, A. (2010). Geographical origin identification of pure Sri Lanka tea infusions with electronic nose, electronic tongue and sensory profile analysis. *Journal of Chemometrics*, 24(3–4), 121–130.
- Moskowitz, H. R. (2001). Sensory Directionals for Pizza: A Deeper Analysis. *Journal of Sensory Studies*, 16, 583–600.

- Pokorny, J., Davidek, J. (1986). Application of hedonic sensory profiles for the characterization of food quality. *Die Nahrung*, 8, 757 - 765.
- Rababah, T. M., Al-u'datt, M., Almajwal, A., Brewer, S., Feng, H., Al-Mahasneh, M., Yang, W. D. (2012). Evaluation of the Nutraceutical, Physiochemical and Sensory Properties of Raisin Jam. *Journal of Food Science*, 77(6), C609–C613.
- Vickers, Z. (1988). Sensory Specific Satiety in Lemonade Using a Just Right Scale for Sweetness. *Journal of Sensory Studies*, 3(15512), 1–8.
- Xiong, R., Meullenet, J. F. (2006). A PLS dummy variable approach to assess the impact of jar attributes on liking. *Food Quality and Preference*, 17(3–4), 188–198.