Barbara Tompos[1]

Received: November 2016 – Accepted: June 2017

Determination of the sensory development directions of beers using the method of penalty analysis

Keywords: product optimization, consumer preference, optimum scale (JAR), beer.

1. Summary

In my research, one of the methods of the determination of sensory product development directions is presented, based on the preference for different beers and the sensory characteristics influencing it. Six different beers were tested, four of which are commercially available (Löwenbrau, Staropramen, cold hop Dreher, Soproni), and two of them are made in breweries (lager, cold hop lager). Beers were rated by 63 college students. 37 of them were women, 26 were men, their age being in the range of 18 to 27 years.

Sensory preference was evaluated by the consumers (naive panelists) on two different scales. First, the evaluation was carried out by characteristics – color, global odor intensity, citrus odour, fruity scent, bitter odour, malt scent, yeast odour, global flavor intensity, citrus flavor, fruity flavor, bitter taste, malt flavor, yeast flavor, sweet taste, sour taste – on a structured optimum scale (just about right, JAR) of 1 to 9 (1=too weak, 5=exactly right, 9=too strong). An important feature of JAR data is that they are bimodal, since not only the deviation from the optimum point, but also the direction of the deviation is important during data processing.

After the evaluation by characteristics, global preference for the products was also provided by testers on a continuously increasing structured scale of 1 to 9 (1=not at all, 2=not very much, 3=moderately not, 4=slightly not, 5=neutral, 6=slightly preferred, 7=moderately preferred, 8=very much preferred, 9=most preferred). The method of penalty analysis has been developed for the combined evaluation of the two scales, and it was carried out by the XL-Stat software. Based on the results of the penalty analysis, it can be determined which sensory properties influence most the global sensory preference for a certain product among consumers, and in what direction it is advisable to change them during product development. Based on the results, judges favored strong, fresh, fruity flavor, and the Dreher cold hop beer was most preferred by far, while the other beers were considered too bitter, unremarkable.

2. Introduction and literature overview

The structure, and the demand and supply sides of the Hungarian food market and, at the same time, of the alcoholic beverage market have changed significantly over the past 25 years. The production of alcoholic beverages was one of the dynamically developing sectors of the Hungarian food industry in the '90s, primarily because of the investments of large multinational companies. The transformation of the demand and supply sides has led to a radical shift of consumer demand towards quality. In addition to changing consumer habits, new product types have also emerged. It is instructive to analyze these changes over a number of years **[1]**.

[1] Szent István University, Faculty of Food Science, Department of Postharvest and Sensory Evaluation

In terms of quantity, household food consumption is measured by the Hungarian Central Statistical Office (KSH) through regular logging performed in the framework of household statistics, and by the compilation of food balances. Food balances in fact mean the amounts of food (supply) available to the population, typically based on feedstocks. Looking at the supply side of alcoholic beverages, the structural transformation of the alcoholic market is clearly evident. On the supply side, the 105.3 liters of beer in 1990 were continuously reduced to 68 liters in 1999 and, after a slight increase, it was reduced further to 66.7 liters by 2007. Wine production was 27.7 liters in 1990, reaching its peak between 2001 and 2006 with 33 to 34 liters, which was followed by a continuous decline, reaching only 21.8 liters in 2014. The volume of distilled spirits decreased from 8.5 liters (1990) to 6.4 liters (2013). Data from the last year show that, while the volume of wine and distilled spirits increased, the volume of wine decreased (**Figure 1**).

Along with the changes in volume, a transformation in the Hungarian beer market could also be observed. Through the liberalization of the beer market that started in the 1990s, in addition to typically foreign-owned large-scale plants, small-scale craft breweries were established, mainly after 2000. The expansion of small-scale craft beer market was primarily promoted by the development of domestic gastro-culture, by the „street food" movement, festivals, competitions and beermaster training courses. The advantage of small breweries compared to large-scale ones is that they are able to adapt to new demands and trends in beer consumption more easily. For the beer produced to meet consumer demand as much as possible, it is essential to map out the needs of the given target group, of the consumer segment, regarding the product. Based on this, the appearance of the beer, its ingredients, the technological steps and the sensory parameters can be adapted to consumer demand in a targeted way **[3]**, **[4]**. At the same time, product selection or the place, time and volume of consumption are strongly socially embedded, it is not possible to apply only the economic approach of supply and demand, since consumer rationality is influenced by several factors **[5]**, **[6]**, **[7]**.

It is recognized by more and more large-scale companies that nowadays it is only possible to produce beer with good organoleptic properties from the highest quality feedstocks, while observing good hygienic practices and adhering to proper technological steps. The main ingredient of beer is water, therefore, the quality of drinking water used in the production is a key factor, which should be analyzed from a microbiological, analytical and sensory point of view as well **[8]**.

Furthermore, in addition to well-established recipes, there is a demand for special beers with a more characteristic taste. Beer consumption habits of the day are increasingly influenced by current trends and new recipes, and these are mainly affected by hops and its newly bred varieties. To enhance the citrus, pine, fruity, spicy aroma of hops, the technological step of cold hopping has gained acceptance. In this method, hops are not boiled, it is only added to the beer in the final stage of hopping, thus it typically gives finished beer a fresh, floral, fruitier, spicier aroma, which can significantly affect the sensory preference for the products. The use of aroma hops has gained acceptance in the Hungarian beer market with the increasing role of small-scale breweries, however, they are also used in large-scale brewing, in accordance with the practices of integrated quality management and food safety systems (Dréher Cold Hop Lager) **[9]**, **[10]**.

Sensory judges are classified by the literature into three categories according to their qualifications: lay/consumer judges, trained judges and expert judges. Different types of task require the use of judges with different qualifications **[11]**.

Trained and expert judges are subject to special, multi-stage judge selection tests, where the "measurement range" of their sensory organs, their thresholds and accuracy are tested, they acquire knowledge related to the field of sensory science, and also gain experience regarding the products to be tested, the scales and the methods **[12]**. Several methods have been developed for the testing of the ability of sensory judges to differentiate, of their repeating ability and of their consensus **[13]**, **[14]**, **[15]**, **[16]**, **[17]**, **[18]**.

Consumer tests can be used to understand which products consumers favor or prefer. Several studies have shown that using only preference (hedonic) measurements it is less possible to predict future preference, therefore, when evaluating the reliability of hedonic methods, the changing behavior of people has to be taken into account **[19]**, **[20]**, **[21]**. The role of consumers (lay judges) in quality management is essential, because the "final control" is exercised by them by buying and consuming the product. Their taste varies widely, it is practically impossible to manufacture products that satisfy the taste of all consumer groups, therefore, it is advisable to optimize the product for the given consumer segment **[22]**, **[23]**, **[24]**.

During food industry product optimization, products are developed according to consumer demand with several aspects in mind. The main aspects are: food safety, nutritional value, sensory characteristics, logistical solutions, etc. From a sensory point of view, several scientific studies have been carried out in the Sensory Analysis Laboratory of Szent István University, including preference tests of aleuron-rich pasta and bread **[25]**, **[26]**, various coffee drinks **[27]**, margarine spreads **[28]**, **[29]**, cola drinks **[30]**, quick-frozen sweetcorn products **[31]**, **[32]**, apple drinks **[33]**, tea drinks **[34]**, flavored kefirs **[35]** and flavored bottled waters **[36]**. The essence of the preference method is that a mathematical relationship is established between the preference test carried out by large number of lay (naive) consumers and the sensory test performed by a small group of expert judges for a given product group **[37]**. Today, new eye tracking methods are known which follow the eye movements of consumers, making product optimization possible **[38**, **[39]**, **[40]**.

Evaluation of sensory tests is usually based on some kind of scale. During scaling, the degree of sensation induced in the brain by the food examined is indirectly analyzed by the judges. Based on the psychophysical approach, if the strength of the physical stimulus increases, then the increase in the sensation can be modeled mathematically **[41]**. Naturally, the possibilities of evaluation are fundamentally influenced by the fact whether a nominal, ordinal, interval or ratio scale is used by the judges. For sensory testing, several types of scales have been developed and used: unstructured line scales, structured line scales, category scales, optimum scales, magnitude estimation scales, etc. **[42]**. Preference tests measure the acceptance of or the preference for a product. In this case, the consumer serves as a measuring instrument, and he or she typically indicates on a scale the value or descriptive term corresponding to their sensation **[43]**, **[44]**. In preference tests, the structured optimum scale (JAR – Just About Right) is often used, in which judges are asked how the given organoleptic characteristics (e.g., sweetness) of the product are perceived: too strong, too weak, or just about right. In the middle of the scale, there is the „just about right" value, while at the two ends are the two extremes for the given property. This test method was created to help researchers understand why consumers like or dislike a product; and how to guide product development in order to increase consumer acceptance **[45]**.

Due to the bipolar nature of optimum scales (JAR), different evaluation methods have been developed. For analysis based solely on the JAR values, the statistical methods of percentage, average, one sample t-test, correspondence analysis, main component analysis, χ^2-test and variance analysis are used **[46]**. The most common evaluation process is penalty analysis, which takes into account preference values as well **[47**, **[48]**, **[49]**.

Of course, several other methods have been published in connection with the modification of penalty analysis, and these are listed according to Gere **[42]**: bootstrapped penalty analysis **[50]**, linear regression **[51],** multiple linear regression **[46]**, multivariate adaptive smoothing curves **[50]**, canonical variate analysis (CVA) **[45]** and the generalized pair-correlation method (GPCM) **[52]**.

3. Objective

As the topic of my research, the preference for different beers and the parameters influencing it were chosen. In the first step, my goal was to determine the development directions of the organoleptic characteristics of different beers. In the second step, I aimed to determine the sensory properties whose change has the greatest impact on the global preference for the product. In accordance with these, I intended to reveal which sensory properties influence the popularity of beers the most, and to what extent judges differentiate between the tastes of premium, non-premium and cold hop products.

4. Materials and methods

4.1. The beers included in the study and beer-making

In the course of our study, six different beers were tested, four of which are commercially available (Löwenbrau, Staropramen, cold hop Dreher, Soproni), while two were produced under pilot-scale conditions, for the Department of Brewing and Distilling of Szent István University (lager, cold hop lager). The raw materials used for making the beers are shown in ***Table 1***.

For the preparation of the beers, the appropriate amounts of malt were first measured using the balance, and then they were ground with the malt mill. The wort tank of the brewing apparatus was filled with 40 liters of water, then the ground malt was added. Following this, warm water was allowed to fill the jacket of the wort tank, and the mixture was heated to 52 °C this way. The malt to water ratio was set to 1:3.5 for the first brewing. The first stage is the protein degradation rest at 52 °C. In this step, endo- and exopeptidases begin to work, causing the protein content of the wort to degrade, which contributes to the increase in foam durability.

The protein degradation step lasted 15 minutes. This was followed by the β-amylase stage (62 °C, 45 minutes). In this step, as an effect of β-amylase, glucose units split off from the non-reducing chain ends in twos, until the enzyme reaches bond α-1,6. The third step is α-amylase degradation (72 °C, 20 minutes). The enzyme produces α-limit dextrins are produced. This step was continued until a negative iodine test was obtained. A sample was taken from the mixing tank with the help of a rod, and it was checked with a drop of iodine whether there was any remaining starch in the wort. After this, the wort was transferred to a filter tub. The wort was introduced through a tube at the bottom of the tub, and air was removed from the tube, so that contact between the wort and the oxygen in air was minimized. After twenty minutes of settling, the liquids were removed. The liquids were transferred from the filter tub to the hop boiler. Water was added to the filter tub, and then solids were washed two more times. This process took 50 minutes. After filtration, the dry matter content of the liquids in weight percentage (Balling degree) was determined. The result of the measurement was 17.7 B°.

Hop boiling lasted 60 minutes. Bitter hops (Perle hops) were added to the boiling cauldron in the tenth minute of boiling. Hops give the bitterness of finished beer. Roughly 10% of the α-acid content of this type of hops is isomerized by heat, and the bitter taste is the result of this transformation. Aroma hops (citra hops) are added to the beer 5 minutes before the completion of boiling. Aroma hops are responsible not for bitterness, but for the fresh citrus taste, its α-acid content does not play a role in the development of the bitter taste. By the end of the boiling, the beer had a Balling degree of 12.9. Following this, the beer rested in a whirlpool for 10 minutes, where the remaining smaller sediments could settle. After this, the beer was cooled in a cooling tank to 12 °C. The cooled beer was transferred to a 13-liter fermentation tub, where 50 g of yeast (Brewmasters lager) was added to it.

The second brew was a cold hop lager, which was made at the same time as the simple lager. The malt types used were the same as in the case of the latter, the only difference being that for 40 liters of water only 10 kg of malt was used. The reason for this was that the initial Balling degree for the simple lager was too high. The initial malt to water ratio here was only 1:4, so the initial Balling degree turned out to be only 15.5 B°. The preparation of this beer was identical to the preparation of the simple lager, with the modification that aroma hops were added to the beer at the very end of the boiling. By the end of the hop boiling the beer had a 10.8 B°. After cooling, the beer was pumped into the 13-liter fermentation tank, where the yeast was added and fermentation began. Fermentation lasted two weeks, and then the beer was drained into bottles. Samples were then transferred to the Sensory Analysis Laboratory of Szent István University for sensory testing.

4.2 Sensory analysis

Beers were rated by 63 college students (37 women, 26 men; age 18 to 27). Consumers (naive panelists) had no previous sensory training, because the goal was not a descriptive analysis of the products, but the exploration of different product development directions, based on consumer acceptance/preference. Sensory preference was evaluated by the consumers (lay judges) on two different scales. First, the evaluation was carried out by characteristics – color, general odor intensity, citrus odour, fruity scent, bitter odour, malt scent, yeast odour, global flavor intensity, citrus flavor, fruity flavor, bitter taste, malt flavor, yeast flavor, sweet taste, sour taste – on a structured optimum scale (JAR) of 1 to 9 (1=too weak, 5=just-about-right, 9=too strong). An important feature of JAR data is that they are bimodal, since not only the deviation from the optimum point, but also the direction of the deviation is important during data processing. After the evaluation by characteristics, overall liking for the products was also provided by testers on a continuously increasing structured scale of 1 to 9 (1=not at all, 2=not very much, 3=moderately not, 4=slightly not, 5=neutral, 6=slightly preferred, 7=moderately preferred, 8=very much preferred, 9=most preferred).

The design and implementation of the sensory experiment, as well as data collection was carried out by a special, cloud-based RedJade sensory software. With its help, random encoding of the beer samples, sample rotation, and computerized execution and data collection of the sensory testing could be performed. The method of penalty analysis was developed for the simultaneous evaluation of the data collected using the two scales, and it was carried out using the XL-Stat software. Following the questions of sensory testing, a short socio-demographic survey was conducted, regarding the sex, age, consumption frequency, place of residence and net monthly income of the testers.

4.2.1 Steps of penalty analysis

Data were evaluated with the XL-Stat software programmed in an Excel environment, using the method of penalty analysis, with the help of the tutorial found on the software's website **[53]**, **[54]**. First, the information of the structured optimum scales of 1 to 9 (JAR) is aggregated by the XL-Stat program into a scale of 1 to 3 for each characteristic (1-4→1=too weak, 5→2=just about right, 6-9→3=too strong). Then the average preference values for the groups „too weak", „just about right" and „too strong" are calculated. Penalties are calculated by subtracting the averages of the „too weak" and „too strong" groups from the average preference value of the „just about right group". The differences obtained are depicted in a so-called „Mean Drop" diagram. In the scatter diagram, differences are plotted against the percentage of consumers in the category. The higher the Mean Drop value and the more consumers perceived the sensory characteristic of the given product too low or too high, the more important the given property (characteristic) is, and it is advisable to adjust it during product development, in order to achieve a higher overall preference for the product among consumers.

On the output side, a large number of tables and charts are generated by program automatically. These are:

1. Descriptive statistics table of the sensory variables examined (number of participants, missing data, minimum, maximum, average, standard deviation values).
2. Spearman's correlation table, showing the correlation between sensory variables. If the correlation proves to be significant (α=0.05), then the variables can be investigated. During the evaluation, mainly the overall liking and the correlation of the JAR variables are important. The question is which JAR variable has a significant effect on global popularity, and how big the effect is.
3. Percentage distribution figure of JAR levels, calculated from the frequency of JAR property values (1-9), and its condensation (1-3).

4. The penalty table includes the following:
   a. JAR dimension names, 3 aggregate levels of JAR data;
   b. frequencies and percentages of each level;
   c. overall liking and average liking of each level;
   d. Mean Drop values for the „too strong" and „too weak" values (these are the differences between the average liking of the JAR levels and the "too strong" and "too weak" levels; this information is important because it shows how many points we lose in liking if a characteristic is „too strong" or „too weak");
   e. the *p*-values, concerning the results of the t-test comparing the averages of the other two levels („too strong" or „too weak") to the average of the JAR level;
   f. automatic interpretation, depending on the level of significance chosen (α=0.05);
   g. weighted differences between the penalty values and the averages (average JAR liking – average liking of the other levels together); (this shows how many points we lose in liking, if the product is not as expected by the consumer);
   h. *p*-values, regarding the comparative test of the JAR level average and the averages of the other levels;
   i. automatic interpretation, depending on the level of significance chosen (here 5%).
5. Mean Drop chart. Too strong (red +) and too weak (blue -) values of the individual properties are displayed. The x axis shows the percentage number of consumers, while the Mean Drop value on the y axis presents the importance of the attribute.
6. Penalty chart. The y axis shows the penalty score, while JAR variables are found on the x axis. The chart shows which variables were significant for each other, which were not, and which variable was characterized by less than 20% of consumers **[54]**.

5. Results

Of the products tested, the Dreher cold hop beer proved to be the most popular, receiving an average liking rating of 6.3. The highest level of consensus among the consumers was also found in the case of this product, since the liking rating had the lowest standard deviation here (1.8). The lowest average liking rating was awarded to the simple beer brewed at the university (4.3), and the cold hop lager beers (4.0). Standard deviation values were also the highest in the case of these beers (2.2-2.4). The liking ratings and standard deviations of the Löwenbrau, Soproni and Staropramen beers were roughly identical (1.9-1.9-2.0) (***Figure 2***).

Interpretation of the results is presented through the example of the Dreher cold hop beer, which received the highest overall liking (OAL) score.

The Spearman correlation table of the Dreher cold hop beer shows the correlation of all variables. From a penalty analysis point of view, the correlations between the overall liking and the other JAR sensory parameters are important. In the interpretation, it should be taken into account that a significant correlation means that the overall liking is less affected by the JAR variable. In an ideal case, correlation is zero (the difference in interpretation comes from the two different scales). Based on the results, there was a significant correlation between several sensory attributes, however, none of the variables affected the liking of the Dreher cold hop beer significantly (***Table 2***).

Figure 3 shows that, for most attributes, the given color, taste/flavour or odor was perceived by the judges mostly to be too weak (blue color) or just about right (green color). What stands out from the values are too weak color intensity, fruity, bitter and malt odour, fruity flavor and too strong sour taste (red color).

The first column of the penalty table of the Dreher cold hop beer (***Table 3***) shows the list of properties evaluated. The second column shows the 3 aggregate levels ("too weak", "just about right", "too strong") created from the original 9 ones. The percentage of the consumers is shown in the third column. In the fourth column, for example, in the first row, the average score given by 50.79% of the judges is shown, i.e., the average popularity value of the consumers who thought that the color of the product was too weak. The fifth column displays the importance of the property endpoints (Mean Drop values). This means that the average of the other two groups is subtracted from the JAR average. Column six shows the result of the *t*-test, indicating whether the differences between the JAR value and the two endpoints are statistically verifiable. The *p* values corresponding to the *t*-test are displayed, and data pairs significant at the 95% significance level are highlighted in bold.

Importance values were calculated and *t*-tests were performed only for the attributes endpoints that were checked by more than 20% of consumers. Column seven indicates the penalty for the entire attributes, i.e., whether the non-conformity of the attribute itself resulted in a decrease in popularity. Column eight shows the result of the *t*-test for the entire attribute, with the *p* values corresponding to the *t*-test, that is, the level at which the result is significant.

The table shows that the null hypothesis can be rejected in the case of the general odor intensity, the fruity, bitter or yeast odour, as well as the global, citrus, fruity, sweet and sour flavor, that is, the averages of the JAR and the other levels belong to different populations. Thus, liking values are significantly influenced by the JAR levels. It follows that, in the case of the global odor intensity, the fruity, bitter or yeast odour, as well as the general, citrus, fruity flavour and sweet taste, because of the too weak level, significantly lower liking scores were given the the product by the consumers. However, in the case of the sour flavour, significantly lower liking scores were given because of the too string level. The same cannot be said about the color, citrus and malt odour, and the bitter, malt and yeast flavor. Here, the null hypothesis is accepted, meaning that liking is not influenced significantly by the levels (***Table 3***).

Figure 4 shows the Mean Drop values (the importance of the attributes) for the two endpoints of the attributes, while the figure below shows the penalties for the attributes as a whole (***Figure 5***). In the case of both diagrams, attributes with no significant differences are indicated in green color, attributes with detectable differences are shown in red, while attributes below the 20% limit, for which the *t*-test is not performed, are colored gray by the program.

Based on the Mean Drop diagram of the Dreher cold hop beer, the most important attributes are global taste and odor intensities, which were found to be too strong and too weak by the consumers, but in the case of the odor intensity, the too weak value is not much lower than the too strong one. Attributes were also considered by the consumers too weak and too strong in the case of the sweet taste and fruity flavor as well. Fruity scent and sour taste were found to be too strong, and bitter odourrs and citrus odourrs too weak by the consumers.

Based on the penalty chart of the Dreher cold hop beer, the greatest penalty was given by the consumers because of the inadequate global flavor intensity and the general odor intensity (***Figure 5***). In addition, fruity scent, bitter odourrs, yeast odourrs, citrus flavor, fruity flavor, sweet taste and sour taste also had a significant effect on liking.

6. Conclusions

Based on the results, it can be concluded that cold hop flavor was most preferred by consumers. This type of beer can be described with fresh, citrus, fruity, less bitter sensory characteristics. In the case of the Dreher cold hop beer, liking was not significantly affected by by either the too low or the too high score of any of the attributes.

Surprisingly, there were hardly any differences between the lager and cold hop lager beers produced at the university. Global flavor and odor intensity were considered too string for both beers by most of the consumers, while citrus and fruity flavor and odour were judged to be too weak and, at the same time, too strong by most. Therefore, they could not differentiate between simple and cold hop flavor. The reason for this could be that, in the case of the beer produced at the university, the flavor characteristic of cold hop beers could not be developed fully. The answer to the question why some of the consumers considered the citrus and fruity character too strong, while other thought it to be too weak might presumably be that a strong flavor and odour was felt by the consumers and its nature could not be decided by them, because they were untrained judges. Overall, it can be stated that these beers could have been more liking had they been produced with less hops, with a less strong flavor.

There were many similarities in the evaluation results of the Löwenbrau, Soproni and Staropramen beers. It can be stated in general that these beer were felt by consumers to be unremarkable, giving them low general flavor and odor intensity scores, they were not popular. The color of Staropramen was considered to be too strong, while that of the other two beers too weak. The flavor of all three beers was judged to be too bitter, therefore, lower preference scores were awarded to them. The similarity of the three beers is an unexpected result, because Staropramen is considered to be a premium beer, while the other two belong to the classic category.

My research shows that penalty analysis is a good and adequate method for the determination of product development directions of foods. Results have supported the fact that consumers prefer fresh, citrus, less bitter flavors, and consider medium, not too strong, but not too weak colors attractive. Of the properties, liking was influenced mainly by general flavor and odor intensity. Premium beer could not be significantly differentiated from the classic category by the consumers surveyed.

In summary, design and implementation of sensory experiments, as well as data collection is adequately supported by the RedJade sensory target software. With its help, random encoding of the beer samples, sample rotation, computerized support of testing and data collection could be performed in an automated and reproducible fashion. The experimental design could easily incorporate the just about right (JAR) scale and questions regarding overall liking. evaluation of the data extracted was efficiently supported by the Excel-based XL-Stat software.

7. Acknowledgement

I would like to thank the staff of Szent István University, Faculty of Food Science, Department of Postharvest Science and Sensory Evaluation: Dr. László Sipos, Dr. Attila Gere and Dr. Zoltán Kókai for their help in the design, implementation and evaluation of the sensory experiments, and also for the help in the statistical evaluation of the RedJade target software and the XL-stat software. I also thank Dr. Szilárd Kun of the Department of Brewing and Distilling for his help in the preparation of the two beers.

8. References

[1] Ernyei, Gy.; Sipos, L. (2006a): Principles of Economics and Management Budapest, BCE Élelmiszertudományi Kar, pp. 174–254.

[2] KSH, STADAT (1990-2014): A rendelkezésre álló alkoholos italok mennyisége.

[3] Ernyei, Gy.; Sipos, L.; Bánáti, D. (2006): Élelmiszerbiztonság és élelmiszerbiztonság menedzsment Szeged, Informen Kiadó, pp. 29–37.

[4] Székely, G.; Sipos, L.; Kiss, O. (2005): Marketing alapismeretek. Budapest, Aula Kiadó, 200 pp. 24–44.

[5] Janky B.; Králik M.; Sipos L. (2005): A fogyasztás társadalmi beágyazottsága. Budapest, BME Kiadó. pp. 7–69.

[6] Sipos, L.; Tóth, A. (2005): A fogyasztói döntés közgazdasági megközelítése. Marketing & Menedzsment, 39 (6) pp. 4–12.

[7] Sipos, L.; Tóth, A. (2006): A közgazdasági értelemben vett irracionális döntések kognitív okai. Marketing & Menedzsment, 40 (1) pp. 22–31.

[8] Sipos, L.; Kovács, Z.; Sági-Kiss, V.; Csiki, T.; Kókai, Z.; Fekete, A.; Héberger, K. (2012a): Discrimination of Mineral Waters by Electronic Tongue, Sensory Evaluation and Chemical Analysis. Food Chem. 135 (4), pp. 2947–2953.

[9] Ernyei, Gy.; Sipos, L. (2006b): Minőségmenedzsment, Quality Management. Budapest, Aula, pp. 23–49.

[10] Jánkfalvi, O. (2016): Aromakomlózott kisüzemi sörök vizsgálata. Szent István Egyetem, szakdolgozat. pp. 5–18.

[11] MSZ ISO 6658:2007 Érzékszervi vizsgálat. Módszertan. Általános útmutató.

[12] Sipos, L. (2009): Ásványvízfogyasztási szokások elemzése és ásványvizek érzékszervi vizsgálata. PhD értekezés. Budapesti Corvinus Egyetem. Döntéstámogató Rendszerek Doktori Iskola. 96-101, 179-184.

[13] Naes, T.; Brockhoff, P. B.; Tomic, O. (2010): Statistics for sensory and consumer science. Wiley, Chicester. pp. 1–287.

[14] Sipos, L.; Kovács, Z.; Szöllösi, D.; Kókai, Z.; Dalmadi, I.; Fekete, A. (2011): Comparison of novel sensory panel performance evaluation techniques with e-nose analysis integration. *Journal of Chemometrics*, 25:(5) pp. 275-286.

[15] Sipos, L.; Gere, A.; Kókai, Z.; Szabó, D. (2012): Application of Artificial Neural Network (ANN) in Praxis of the sensory evaluation: Mesterséges ideghálózatok (ANN) alkalmazása az érzékszervi minősítés gyakorlatában. Élelmiszervizsgálati Közlemények 58:(1-2) pp. 32-46.

[16] Sipos, L.; Gere, A.; Szabó, A.; Kovács, S; Kókai, Z. (2013): Multivariate Methods For Assessing Sensory Panel Performance. In: Héberger K (szerk.) Programme and Book of Abstracts of CC 2013 - Conferentia Chemometrica. Konferencia helye, ideje: Sopron, Magyarország, 2013.09.08-2013.09.11. Budapest: Hungarian Chemical Society, p. 6.

[17] Gere, A.; Ladányi, M.; Dürrschmid, K.; Sipos, L. (2014): Comparison of the computation method of panellist's discrimination ability, applied in MAM-CAP, to other approaches In: Hydling Grehte, Nielsen Jette (szerk.) Eurosense: A Sense of Life. Konferencia helye, ideje: Copenhagen, Dánia, 2014.09.07-2014.09.10. Oxford: Elsevier, 2014. p. P136.

[18] MSZ ISO 11132:2013 Érzékszervi vizsgálatok. Módszertan. Általános irányelvek a leíró vizsgálatot végző bírálóbizottság teljesítményének mérése

[19] Mahesvaran, D. (1994): Country of Origin as a Stereotype: Effects of Consumer Expertise and Attribute Strenght on Product Evaluations, The journal of consumer research, 21, 2 354–365.

[20] Székely, G.; Sipos, L.; Kiss, O. Zs.; Kocsis, M. (2006a): Basic Marketing Budapest, Aula Kiadó, pp. 64–99.

[21] Székely, G.; Sipos, L.; Losó, V. (2006b): FMCG marketing. Budapest, Aula Kiadó, 2009. pp. 130–147.

[22] Kókai, Z.; Sipos, L. (2011): Érzékszervi minősítés. Budapest, Nemzeti Tankönyvkiadó.

[23] Sipos, L.; Soós, R.; Pádár, K. (2007): Ásványvíz-preferenciák a fiatalok körében. Marketing és menedzsment, 41:(3) pp. 39–51.

[24] Sipos, L. (2008): A conjoint elemzés mint a vásárlói preferenciák vizsgálatának eszköze az élelmiszergazdasági marketingben. Marketing & Menedzsment, 42 (3) pp. 4-13.

[25] Bagdi, A.; Szabó, F.; Gere, A.; Kókai, Z.; Sipos, L.; Tömösközi, S. (2014): Effect of Aleurone-Rich Flour on Composition, Cooking, Textural, and Sensory Properties of Pasta. LWT - Food Science and Technology. pp. 996–1002.

[26] Bagdi, A.; Tóth, B.; Lőrincz, R.; Szendi, S.; Gere, A.; Kókai, Z.; Sipos, L.; Tömösközi, S. (2016): Effect of Aleurone-Rich Flour on Composition, Baking, Textural, and Sensory Properties of Bread. Lwt-Food Science and Technology. pp. 762–769.

[27] Várvölgyi, E.; Gere, A.; Szöllősi, D.; Sipos, L.; Kovács, Z.; Kókai, Z.; Csóka, M.; Mednyánszky, Z.; Fekete, A.; Korány, K. (2014): Application of Sensory Assessment, Electronic Tongue and GC–MS to Characterize Coffee Samples. Arabian Journal for Science and Engineering. pp. 125-133.

[28] Györey, A.; Gere, A.; Kókai, Z.; Molnár, P.; Sipos, L. (2012a): Effect of Sample Presentation Protocols on the Performance of a Margarine Expert Panel. Acta Alimentaria. pp. 62–72.

[29] Györey, A.; Gere, A.; Kókai, Z.; Sipos, L.; Molnár, P. (2012b): Kenőmargarinok Bírálatára Kiképzett Szakértői Panel Teljesítményének Mérése. Élelmiszervizsgálati közlemények - Journal of Food Investigations. pp. 47–58.

[30] Szöllősi, D.; Kovács, Z.; Gere, A.; Sipos, L.; Kokai, Z.; Fekete, A. (2012): Sweetener Recognition and Taste Prediction of Coke Drinks by Electronic Tongue. Sensors Journal, IEEE. November, pp. 3119-3123.

[31] Gere, A.; Losó, V.; Györey, A.; Kovács, S.; Huzsvai, L.; Nábrádi, A.; Kókai, Z.; 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. pp. 3213–3225.

[32] Losó, V.; Gere, A.; Györey, A.; Kókai, Z.; Sipos, L. (2012a): Comparison of the Performance of a Trained and an Untrained Sensory Panel on Sweetcorn Varieties with the Panelcheck Software. Applied Studies in Agribusiness and Commerce – APSTRACT. pp. 77–83.

[33] Kovács, Z.; Dalmadi, I.; Lukács, L.; Sipos, L.; Szántai-Kőhegyi, K.; Kókai, Z.; Fekete, A. (2010): Geographical Origin Identification of Pure Sri Lanka Tea Infusions with Electronic Nose, Electronic Tongue and Sensory Profile Analysis. J. Chemom. 24 (3–4), pp. 121–130.

[34] Kovács, Z.; Sipos, L.; Szöllősi, D.; Kókai, Z.; Székely, G.; Fekete, A. (2011): Electronic Tongue and Sensory Evaluation for Sensing Apple Juice Taste Attributes. Sens. Lett. 9 (4), pp. 1273–1281.

[35] Gere, A.; Kovács, S.; Pásztor-Huszár, K.; Kókai, Z.; Sipos, L. (2014b): Comparison of Preference Mapping Methods: A Case Study on Flavored Kefirs. Journal of Chemometrics. pp. 293–300.

[36] Sipos, L.; Gere, A.; Szöllősi, D.; Kovács, Z.; Kókai, Z.; Fekete, A. (2013): Sensory Evaluation and Electronic Tongue for Sensing Flavored Mineral Water Taste Attributes. Journal of Food Science. pp. S1602–S1608.

[37] Losó, V.; Tóth, A.; Gere, A.; Heszberger, J.; Székely, G.; Kókai, Z.; Sipos, L. (2012b): Methodology Problems of the Industrial Preference Mapping. Acta Alimentaria. pp. 109–119.

[38] Danner, L.; Antoni, N.; Gere, A.; Sipos, L.; Kovács, S; Dürrschmid, K. (2016): Make a Choice! Visual Attention and Choice Behavior in Multialternative Choice Situations, Investigating Different Food Product Groups. Acta Alimentaria, 1. pp. 1–5.

[39] Gere, A.; Danner, L.; Antoni de N.; Kovács, S.; Dürrschmid, K.; Sipos, L. (2016): Visual attention accompanying food decision process: an alternative approach to choose the best model. Food Quality and Preference, 51. pp. 1–7.

[40] Kovács, E.; Gere, A.; Székely, D.; Kókai, Z.; Sipos, L. (2016): Szemkamerás Vizsgálatok Egy élelmiszer Fogyasztói Megítélésében. Élelmiszervizsgálati közlemények – J. Food Investig. 62 (2) pp. 1048–1061.

[41] Lawless, H. T.; Heymann, H. (2010): Sensory Evaluation of Food, 2st ed. Chapman and Hall, New York, NY. 243–246.

[42] Gere, A. (2016): Módszerfejlesztés a preferencia-térképezésben. PhD értekezés, Szent István Egyetem. 7–47.

[43] Meilgaard, M.; Civille, G. V.; Carr, B. T. (1999): Sensory Evaluation Techniques, New York: CRC Press. pp. 5–387.

[44] Stone, H.; J. Sidel. (1985): Sensory Evaluation Practices. San Diego, CA: Academic Press. 1993. Sensory Evaluation Practices, 2nd ed. San Diego, CA: Academic Press, Inc.

[45] Popper, R.; Gibes, K. (2004). Workshop summary: Data analysis workshop: getting the most out of just-about-right data - Abstracts. Food Quality and Preference, 15 (7-8), 891–899.

[46] Pagès, J.; Berthelo, S.; Brossier, M.; Gourret, D. (2013): Statistical penalty analysis. Food Quality and Preference, 32, pp. 16–23.

[47] Plaehn, D. (2013): What's the real penalty in penalty analysis? Food Quality and Preference 28, 456-469.

[48] Popper R. (2004): Workshop summary: data analysis workshop: getting the most out of just-about right data, Elsevier Science.

[49] ASTM MNL-63. (2009): ASTM MNL-63 Just-About-Right (JAR) Scales: Design, Usage, Benefits, and Risks.

[50] Meullenet, J-F.; Xiong, R.; Findlay, C. F. (2007): Multivariate and Probabilistic Analyses of Sensory Science Problems. Wiley-Blackwell, New York, NY. pp. 27–47.

[51] Plaehn, D.; Horne, J. (2011): A Regression-Based Approach for Testing Significance of JAR Variable Penalties. (L. Rothman és M. J. Parker, Szerk.) Just-About-Right.

[52] 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, pp. 88–96.

[53] XL-Stat 2012.6.2 manual. (Addinsoft, 28 West 27th Street, Suite 503, New York, NY 10001, USA).

[54] Penalty analisys in Excel tutoial, 2016. https://help.xlstat.com/customer/en/portal/articles/2062259-running-a-penalty-analysis-with-xlstat?b_id=9283 (Aquired: 16.01.2018).