H. Scheuren, J. Menze and K. Sommer

A New Approach to Sensory Evaluation

Flavour is the significant criterion of evaluation in beer tasting, which depends on many primary and secondary causes. This includes desirable and un-desirable aromatic compounds, but also haptic and physical issues like viscosity, the CO_2-level and colour. A beverage for example with a nontypical colour presented in a wrong drinking vessel is always evaluated differently to the "correct" form without any modifications.

Next to these problems taste in general is a subjective matter and is valued individually quite different. Furthermore the influences of process changes on taste are difficult to evaluate. In order to get representative and fast results in the brewing and beverage technology, a sensory test developed by Sommer was used [1]. The test consists of a dual testing, therefore one sample is the product with the original taste and the second one is the product with a changed aromatic profile. The proband has to detect the changed sample. For the case he is not able to determine it, he has to guess. This test is repeated several times with a changing taste difference of both samples. New about this procedural is the interpretation of the testing results by using a probabilistic evaluation form. Thereby an intensity curve is given which indicates the relation between rising taste difference and consumer taste perception.

The needed requirements are the same as for every taste test. The relative number of test persons, who can detect a difference in change, remains almost constant in a group and more or less independent from external influences. Individual errors follow statistic behaviour and can be assessed if the number of tasters is big enough. The significance of results from a smaller group of experienced tasters in a sensory panel can be improved by increasing the number of test persons even if they are less experienced.

The results of this work are based on differentiation trials with beer of different concentrations of Benzaldehyde as a typical off-flavour compound [2]. The usage of this compound is caused by its concentration depended nontoxity and the well known marzipan taste. In a group of inexperienced tasters, mainly young food technology students, qualified differentiation showed high conformance and significance. This test indicates that the method with the incorporation of statistical methods can be used efficiently for the evaluation of process changes and their effect on beer taste. For the performed Benzaldehyde beer mixing process the developed functional coherence can be applied for economic process optimisation. So for the possibility of creating a new and atypical marzipan beer which can taste 99 % of the consumers the needed dosage of Benzaldehyde could be calculated.

Descriptors: sensory evaluation, taste, flavour, Benzaldehyde

1 Introduction

In the Brewing industry, as in the life science industry in general, taste testing is a very effective tool for describing and controlling the quality of food as for example beer. There is a broad range of different taste tests like a duo test [3], a triangle test [4], a duo triangle test [5], a test called "A" and "not A" testing [6] and a ranking test [7]. All these methods belong to the group of discrimination tests which have a quantification of tested samples in common. In contrary to them are the descriptive test methods in which a quality analysis of the samples is the purpose [8]. Therefore a defined vocabulary and special graphical presenting technique like the flavour wheel are used. Additional and detailed information can be found in the cited literature.

In brewery there are two principal testing tools. One form consists in the usage of an expert panel which can detect small differences in the sensory quality. Such a group is very helpful for an engineer concerning the finding and detecting problems in the production process. To receive an expression about the consumer acceptance, this tool is useless, because a trained tester tastes more successfully than an untrained one.

Another form of taste testing is panels which consist of a high number of consumers. The sense of taste of these persons is mostly untrained and the resulting statements about the product quality are dependent to personal preferences. With this way of evaluation the producer cannot optimize the production process, but he can get information about the consumer's behaviour.

The following method can be seen as a combination of both kinds of tests. Therefore the results of a high numbered tasting panel are evaluated by using statistical considerations. By that way it is possible to calculate and to predict changes in the aromatic profile of beer.

2 Basics

In the used test every proband has to taste and to compare two samples in a special number of runs. One sample is untreated beer, the other one is a beer flavoured with a changing concen-

Authors:

Dipl.-Ing. Hans Scheuren, B.Sc. Julia Menze, Univ.-Prof. Dr. Ing. Karl Sommer, Technische Universität München, Lehrstuhl für Verfahrenstechnik disperser Systeme, Maximus-von-Imhof-Forum 2, 85354 Freising-Weihenstephan, Germany, e-mail: H.Scheuren@lrz.tum.de

Tables and figures see Appendix

tration of an aromatic component. The test person has to detect the treated beer dependent to his personal tasting ability. So an untrained tester cannot detect the aromatised beer sample if the concentration difference between both samples is very low. Instead a trained proband can determine the treated beer sample in this concentration range.

A very important element of the test consists in the force for decision. If an untrained taster cannot detect a difference between two samples he still has to declare one beer as flavoured. By that way the number of all persons is divided to both samples in two groups. The only possibility to influence the number of persons in every group consists in the raising or decreasing of the content of the aromatic component in the treated beer. A higher concentration of an aromatic component leads to more probands who are able do detect the treated sample. The result is a distribution in the exemplary form as it is given in figure 1.

The abscissa describes the exemplary concentration of a not defined exemplary aromatic component (*I*); the ordinate illustrates the absolute number of people (*N*) who detect the sample with a special concentration of the used flavour.

If two samples are identically (the aromatic profile is the same, no beer is treated) no proband is able to detect a difference. So caused by the force to decide, all probands are still divided in two groups. In the scheme this case can be recognised. For an exemplary aromatic concentration of $I_0 = 0$ a number of $N_0 = 50$ testers declare one of the two samples as treated. The difference to the over all number declares the other sample as treated. It is obvious that with a raising concentration of an aromatic component in one of both samples a raising number of persons can detect that difference. At the end a very high concentration of an aromatic substance in beer is detected by 100 of 100 people. So every proband can taste the difference between both samples.

The interesting question is now, how can the number of correctly testing probands be calculated. The answer to that question is given by statistical or probability considerations. In a run every proband has to decide for a sample. Sometimes he tastes a difference, sometimes not. In the case of two identical samples every trained or untrained proband has to guess. And to guess means to have a success possibility of 50 % to choose the apparent right one. Without any deviation from the statistical distribution 50 of 100 probands declare the first sample as flavoured and the other 50 people decide for the second one.

With an increasing aromatic concentration the number of guessing people is decreasing. By knowing the guessing probability of 50 % it is possible to calculate the number of people which are in fact able to determine a difference in the taste of both beer samples. Therefore the following equations have to be used.

3 Calculation Method

A describes test panel of 100 people is signed as *N*. That means:

$$N = 100 \tag{1}$$

In every test run all these persons have to evaluate two samples and they have to decide which of both samples is the one containing a concentration of the unwanted aromatic component. As a result of this the amount N is divided in probands who give the correct answer (*A*) and the persons who give the wrong answer (*B*). Formulated in an equation brings:

$$N = A + B \tag{2}$$

Considering the described problem of the statistical relation for interpreting the test result we have to understand that the number of persons who determine the right sample must be divided in the ones who taste it correctly (A^*) and the ones who do not taste but choose correctly (A'). In an equation it means:

$$A = A^* + A' \tag{3}$$

Equation 3 inserted in equation 2 leads to:

$$N = A^* + A' + B \tag{4}$$

The taste test shall give an answer to the question which number of probands or which percentage of them tastes the difference. That means we are looking for the value of the parameter A^*. By using the following equation we get the number of the correctly tasting probands A^*:

$$A^* = N - B - A' \tag{5}$$

Solving this equation is not possible because we do not have any information concerning the parameter A'. So we do not know how many probands are guessing correctly.

Useful information for solving this problem gives the parameter *B*. This factor describes the amount of probands who do not taste a difference between both samples. If they cannot detect the correct sample but have to decide for a sample, they must guess. The chance for guessing sample A or sample B is 0.5. The number of persons who guess wrong is the same as the number of the ones who guess correctly:

$$A' = B \tag{6}$$

Inserting equation 6 in equation 5 leads to:

$$A^* = N - 2 \cdot B \tag{7}$$

The number of correctly tasting persons can be calculated by equation 7. Scaling this result by the over all number *N* leads to a dimensionless expression.

4 Material and method

The idea of the test is to determine a distribution between a changing concentration of an aromatic component in beer and the ability of consumers for detecting the treated samples. Using the described formulas it must be possible to calculate the probands who are really able to detect a special concentration.

As aromatic component the substance Benzaldehyde is chosen, because the taste of it can be simply described and communicated as marzipan like, furthermore it is an aging indicator for beer. Independent from the testing method the results should be interesting for the brewing industry. Also in the used concentration range Benzaldehyde is nontoxic and not dangerous. At least the handling concerning the dosing and mixing is compared to e.g. Dimethylsulphide caused by the low steam pressure advantageous. The used medium is beer flavoured with different amounts of Benzaldehyde. This substance has a taste threshold in beer of 2 $\mu l/l$ [9]. The tested concentrations vary over a range from 0 to 5 $\mu l/l$;

The described method is tested at the Technische Universität München (TUM) and consists of duo tests. Therein every proband tests and compares six times two samples A und B with different concentrations of Benzaldehyde. The test term is about 3 days. The results of the different days are summarised to a scope of testing, so the over all number of tested persons is about 273 students.

The resulting data are presented in two figures. Figure 2 contains the counted amount of testers who declare sample A as the treated sample. In the next figure the relative number of the correctly tasters, using the relative form of equation 7, is presented.

5 Results

The following figure shows the result. The abscissa describes the concentration of Benzaldehyde, the ordinate illustrates the absolute number of people who detect the flavoured sample.

The experimental values confirm the theory. For the case of two identical beer samples ($I_0 = 0\ \mu l/l$), every proband has to guess with a guessing probability of 50 %. So nearly the half of 273 testers declare sample A for the apparently treated one, whereas the rest of all probands declare sample B as the beer with a special concentration of Benzaldehyde.

Low concentrations of Benzaldehyde under the threshold ($I \leq 2\ \mu l/l$) are difficult to taste and most of the probands still have to guess. The guessing probability is unchanged so nearly the half of the test panel decides for sample A and the rest for sample B. But only a few probands choose sample A as the treated one because they really can taste the difference to sample B. Caused by these few testers the slope of the distribution curve is changed.

For higher concentrations of Benzaldehyde ($I \geq 2\ \mu l/l$) the number of correctly tasters who really taste the difference between both samples is rising. At the highest concentration of the test series nearly everybody can apparently taste the treated sample.

Answering the quantitative question how many testers are able to really taste a difference between the flavored and the unflavored beer the relative form of equation 7 is used. The following figure compares the known experimental tasting curve (red) of figure 2 with the calculated one (blue).

The difference between the red curve and the blue one is obvious. For the case of two identical samples ($I_0 = 0\ \mu l/l$) the guessing probability is not confirmed.

Another interesting result can be seen for the highest concentration $I_5 = 5\ \mu l/l$. Reading the red curve it seems that nearly every proband is able to detect the difference. Calculating the amount of correctly testers a difference can be seen. So a concentration of Benzaldehyde which is more than two times higher than the threshold of this component in beer cannot be detected by all probands.

Table one contains the test results and the calculated number of persons, who taste correctly.

Repeating the example of an aromatic concentration of 6 $\mu l/l$ the flavored sample A is detected by 248 persons correctly. Using the described equations the number of probands who really can taste the difference is 196 (of a total number of 273 probands).

For a zero concentration of Benzaldehyde a number of 132 persons give the correct answer. Calculating the number of correctly tasting people the result is –9. The negative number results by the calculation and can be explained by statistical derivations. Furthermore the test procedure itself contains a problem. At the beginning of every test series the proband needs a guide line of decision, for example the order:

“Detect the one of two beer samples which has a non beer typical taste”

That simple sentence causes a problem that way, that there is no beer typical taste. Furthermore in the used concentrations of Benzaldehyde some probands sense the flavored beer sample as the one with typical beer taste. The sweet, almond past taste of Benzaldehyde “improves“ for some probands in low concentration the taste of beer, so the untreated beer sample seems to have the worse taste.

A first solution can be found in modifying the guide line by additional information, for example:

“Detect the one of two beer samples which has a non beer typical, marzipan like taste”

This guide line contains a better description of the nontypical taste. For higher concentration over the threshold of Benzaldehyde every proband has a correct deciding line. But for low concentrations the problem still exists because unrecognized and in combination with other aromatic component Benzaldehyde can still possibly influence the over all taste.

A second solution for this problem consists in expanding the number of samples. Comparing the taste of three samples allows avoiding the description of the non beer taste. Therefore a new guide line has to be given. For example:

“Detect the one of three beer samples which has a divergent taste”

An improved quality of the calculated number of correctly testing probands has to be expected. That triangle test will be published in a further article and the distribution curves of both tests will be compared.

6 Summary

The innovation of the presented test consists in the determining of a distribution between a changing aromatic beer profile and the ability of a consumer panel for detecting it. The results determine the presented method and the idea behind. Furthermore the threshold of Benzaldehyde in beer is firmed by the distribution.

The advantage of the described method consists in its simplicity. For performing the test and getting results no expert or trained testers are needed. An interesting application can be seen in the use of this method as a simple analysing tool. For example the impact of a new aggregate in the brewhouse can be quantified in a single test by asking a staff of consumers. If only a low number of these probands can test a difference between a traditionally produced beer and a new one (e.g. a relative part of 1 %), than the product beer could be judged as "unchanged". So the described testing method could be used by brewers and brew house constructers to quantify the influence of new aggregates in a traditional brew process.

In a next step the presented duo test method will be compared with a triangle test method based on the same probability considerations.

7 Literature

1. Sommer, K.: Geschmackstest zur Ermittlung der Qualitätsfunktion an Beispiel von Schokolade. In: Fette, Seifen Anstrichmittel **78** (1976) 4, pp. 176-178.
2. Narziß, L.: Abriss der Bierbrauerei, 6., neu bearb. Aufl. Weinheim 2001.
3. Deutsches Institut für Normung: Sensorische Prüfverfahren – Paarweise Vergleichsprüfung (1997) DIN 10954. Berlin.
4. Deutsches Institut für Normung: Sensorische Analyse – Prüfverfahren -Dreiecksprüfung (2005) DIN EN ISO 4120. Berlin.
5. Deutsches Institut für Normung: Sensorische Prüfverfahren – Duo-Trio-Prüfung (2003) DIN 10971. Berlin.
6. Deutsches Institut für Normung: Sensorische Prüfverfahren –„A“-„nicht A“-Prüfung (2003) DIN 10972. Berlin.
7. Deutsches Institut für Normung: Sensorische Prüfverfahren – Rangordnungsprüfung (1997) DIN 10963. Berlin.
8. Heyse, K.-U. (Hrsg.): Praxishandbuch der Brauerei. Hamburg 2000.
9. Meilgaard, M. C.: Flavor Chemistry of Beer: Part 2: Flavor and Treshhold of 239 Volatiles: MBAA Technical Quaterly.

Received 20 September, 2010, accepted 27 November, 2010

Appendix

Table 1 **comparison of experimental and calculated values**

Aromatic concentration [µl/l]	test results	calculated results
0	132	–9
1	135	–3
2	170	67
3	213	153
4	229	185
5	248	223

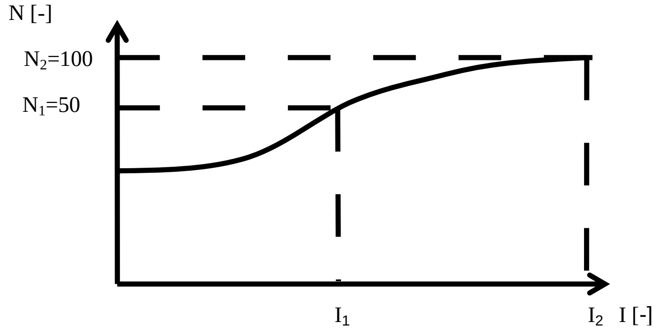

Fig. 1 **Distribution of 100 probands (N) related to changing concentrations of an aromatic component (I)**

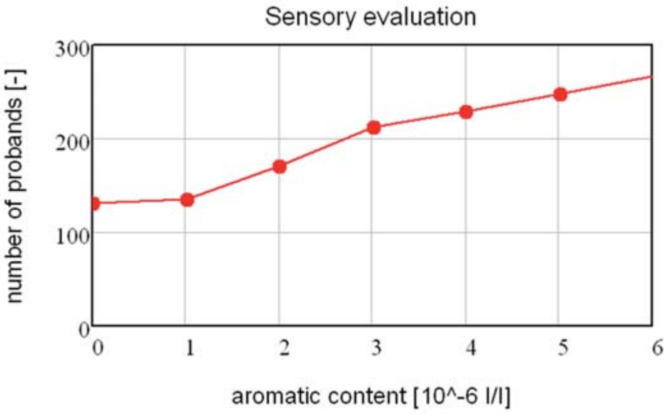

Fig. 2 **Results of the taste test**

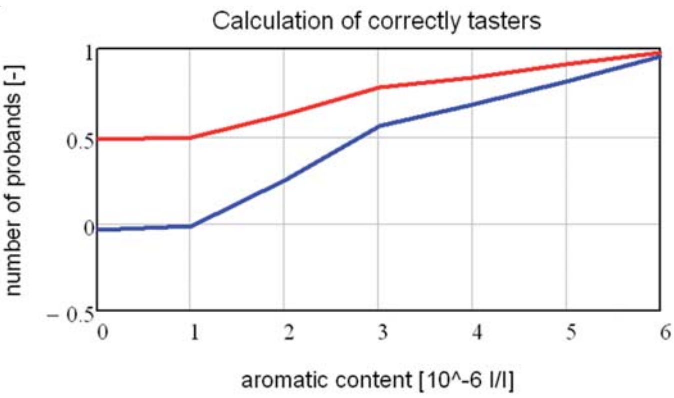

Fig. 3 **Calculation of the correctly tasters**