Food Sci. Biotechnol. Vol. 16, No. 6, pp. 863 ~ 867 (2007)

MINI REVIEW

Food Science and Biotechnology

Some New Approaches to Consumer Acceptance Measurement as a Guide to Marketing

Hye-Seong Lee* and Michael O'Mahony[1]

Department of Food Science and Technology, Ewha Womans University, Seoul 120-750, Korea
[1]Department of Food Science and Technology, University of California, Davis, CA 95616, USA

Abstract The potential impact of the methods of sensory science on consumer testing and marketing is reviewed. Areas such as predicting purchase behavior, new approaches to scaling, and cross cultural effects are discussed. An example of the complexity of sensory measurement used to obtain consumer and marketing information is highlighted, using the simple paired preference test as an example.

Keywords: marketing, consumer behavior, consumer testing, sensory science, paired preference test

Introduction

Consumer/sensory science will not increase a product's performance in the marketplace; that is the responsibility of the marketing department. However, accurate data from consumer/sensory science can provide accurate and valid measures of consumer acceptance and intent to assist in the decision making process. How much will consumers like the taste of a new product; will they want to buy it? What combination of attributes would be the most attractive to consumers? Has the right target market been chosen? How can market share be protected? What packaging would attract the most potential buyers? Which marketing/advertising campaign will be the most persuasive? There are many questions which come by a marketing department and often decisions are made without any regard to the careful collection of unbiased, valid information. The consumer/sensory scientist can gather the facts to guide the decision makers in the company, in a way that goes far further than simple marketing research.

Consumer/sensory science is really an applied branch of experimental psychology. Sensory scientists measure human behavior. Rather than guessing what people like or wish to buy, they make controlled measurements to predict it. Yet, such measurement is complex and requires highly trained professionals. It has long since passed the point where the president and a few trusted employees decide what is marketable. The science has come a long way further than simple taste testing in the kitchen.

In America, some more adventurous companies are realizing that there is no strict delineation between sensory evaluation, consumer testing and marketing research. They are all trying to predict and manipulate the outcomes of the same thing: human behavior in the marketplace. So they are being combined into multidisciplinary teams which are only just beginning to realize the power that they get from consumer/sensory measurement. The typical marketing person uses inspiration or his experience to make predictions. The consumer scientist often uses questionnaires and focus groups to try to predict sales. But the sensory scientist has many techniques derived from the science of the measurement of human behavior and is now beginning to introduce them into the area of measuring consumers' reactions and possible market outcomes. It might not seem that an analysis of sensory traffic in the nervous system, using the jargon of communications engineers like signal-to-noise ratios which are concerned in Thurstonian approach for sensory tests (1, 2), can help in marketing, but it can.

There are many sophisticated computer programs for making decisions regarding the best type of product to market, for example Preference Mapping (3) and Landscape Segmentation Analysis (4). Computing and statistics are well advanced. Most of such methods use multivariate reduction techniques creating only a few new variables from the original, multi-dependent variables, without loosing much information. Principal component analysis (PCA), which is commonly used for summarizing the relationship between various products and their sensory attributes in food science (5-7), is an example of such multivariate analysis. Yet, what is not so advanced is the development of efficient, valid, and bias free techniques of behavioral measurement. Most of the techniques used in this area were developed around 50 years ago and much more is now known about the sensory system and decision making processes in the brain. The influx of sensory scientists, with their research skills at developing new experimental techniques has caused advances in measurement techniques, which will be useful for professionals involved in marketing and consumer testing.

There are too many problems in this area for the scope of this article. Some will be mentioned briefly while one area will be considered in more detail as an example of the complexity involved.

Predicting Purchase Behavior

One obvious measure of interest to the marketer is purchase intent. How will the product perform in the

*Corresponding author: Tel: +82-2-3277-3624; Fax: +82-2-3277-4213
E-mail: hlee@ewha.ac.kr
Received August 22, 2007; accepted September 11, 2007

market place? Who will buy it and how often? The most valid way to measure this is to test-market the product and monitor how often the consumers buy it. The technology is available and demographic information can be obtained from shop loyalty cards. Yet, if a researcher does not have access to such 'in house' behavioral data, the alternative is to ask the consumers whether they like the product or would buy it, using some method of acceptance testing, and hope that consumers' responses will predict their behavior. The consumer/sensory scientist will want any such method to be validated by monitoring the consumer's purchases after testing, to determine whether purchase behavior of the consumers was correctly predicted by the test. Yet, because this is a lengthy task, it is rarely done. Rosas-Nexticapa *et al.* (8) monitored judges' purchases of yogurts, using home visits for a year, after requiring consumers to rate how much they liked the products and their purchase intent. Consumers were required to save all their food receipts and were visited every two weeks, under the guise of receiving nutritional advice for their families. Comparing purchases with consumers' ratings, it was seen that the yogurt that received the highest rating on the scale was the one that consumers purchased throughout the year. There was no predictive association for the products that were rated second and third. The information obtained was no better than asking consumers which product they liked the best or they were most likely to buy. Further, for the purchase intent scale, this prediction was only true for a mere 27% of the consumers (68% allowing ties with other yogurts). This is the only such study but it begins to provide marketers with some idea of the predictive validity of this technique.

New Approaches to Scaling

Despite its only small predictive success in the market-place, consumers are often asked to rate products for liking using some form of numerical value. The most commonly used scale is the 9-point hedonic scale developed by the US Army Quartermaster Food and Container Institute, for predicting soldiers' food choices in their canteens as far back as 1949. The scale was described by Peryam and Girardot (9), Peryam and Pilgrim (10) and Peryam *et al.* (11). It was based on developmental work by Jones and Thurstone (12) and Jones *et al.* (13), using scaling techniques developed by Edwards (14). Again, the sensory consumer/scientist would not be satisfied with using such a scale until the cognitive decision processes associated with its use were understood. Various lines of research: effects of forgetting during scaling (15-18), context and range biases (19-22) have indicated that when untrained consumers use such scales, they are using a 'relative' decision rule. In other words, they are merely ranking the products and using the scores on the scale to represent the perceived distance between the ranks.

Pipatsattayanuwong *et al.* (23) were asked to investigate at what temperature consumers wished to drink their coffee and at what temperature they expected to be given coffee to drink. This was after many legal cases had been brought in the USA, regarding burns caused by coffee spills. To make the testing experience as simple as possible for consumers, it was decided that they should be tested using as simple a method as possible and one with which they were familiar: ranking. The spacing between the ranks was computed using a signal detection analysis and was represented by a nonparametric measure of preference, the R-Index (24, 25). This is essentially a probability value. In this context, it represents the probability of one product being preferred to another in a pair-wise comparison. A value of 50% represents one product having an equal chance of preference over the other (no preference, zero distance between the ranks). A value of 75% indicates a 75:25 split (a bigger spacing) and 100% represents complete preference (very large spacing). The advantage of this approach is that the spacing between the ranks is based on consumer behavior (their rankings) and not on their unskilled numerical estimates of degree of liking.

Lee and O'Mahony (26) went on to apply this technique to measurement to a marketing problem of measuring an abstract consumer concept. The concept was the appearance of refreshingness of toothpastes when spread on the toothbrush. Did a blue toothpaste appear more refreshing than a pink or striped one? Here again the task for the consumer was simple; it was merely to rank a set of twenty toothpastes for their appearance of refreshingness. R-Indices were used to describe the spacing between the mean ranks. They were also used to give numerical values for how close to the centre of the refreshingness concept each toothpaste was. Lee and O'Mahony (26) went on to demonstrate that computer generated pictures could be substituted for actual toothpastes, making the testing of product ideas much faster and simpler.

Cross-Cultural Effects

As world food markets become globalized (27, 28) companies need to market products that will gain consumer acceptance in a variety of cultural settings, which highlights the need for suitable methods for cross-cultural sensory experimentation (29). One such focus is the area of cross-cultural measurement of liking. Yeh *et al.* (30) noted that Korean, Chinese, and Thai consumers used a smaller range of scores on a 9-point liking sale than Americans. They hypothesized that these East Asian consumers, because of their culture of politeness, were reluctant to express negative responses. Yet, Prescott *et al.* (31) did not find this effect with young urban Japanese. Further research (32) indicated, once again, that compared with Americans, Koreans, and Japanese used a smaller range of scores. Yet, the reasons were different. The Japanese tended to feel that it is impolite to use extremes and their use of the scale categories gradually dropped towards the ends of the scale. Koreans, on the other hand, were reluctant to use just the ends of the scale, which were labeled 'like extremely' and 'dislike extremely'. It appeared that the term 'extremely' has a much more extreme meaning in Korean than in English. For a global company, the tendency of consumers to use different ranges of the scale causes problems for comparison of consumers' responses among different countries. It has initiated much research effort. Yet, the solution is simple. Requiring consumers to simply rank the products under consideration automatically ensures the range of responses is the same in every country. As above, R-Index measures can be used to determine the spacing between the ranks.

Paired Preference; An Example of Complexity

To gain an idea of the complexity of consumer/sensory measurement, it is worthwhile looking in detail at an apparently simple form of measurement: the paired preference test (33-35). This is a common test both in marketing and consumer measurement to determine consumer preference. The consumer is simply asked to state which of two products she prefers or whether she has no preference. It should be straightforward.

'No Preference' option Yet, the first disagreement comes with the argument about the use of the 'No Preference' option. Without this option, a simple binomial statistical analysis is possible. It is argued that if one probes hard enough, a consumer will eventually reveal a deep seated preference in her psyche. Yet, if a preference is so deep seated, it is unlikely to be operating in the buying situation where any preferences would need to be 'near the surface'. Yet, to allow a simple statistical analysis, some researchers still choose not allow such an option or simply ignore it. There are various suggested ways of eliminating 'No Preference' data from the statistical analysis. These have been reviewed but not necessarily recommended (34-39). These range from simply ignoring the 'No Preference' responses, splitting them equally between the two preference options or splitting them proportionately between the preference options according to the respective preference frequencies. A further rather bizarre option is to assign the 'No Preference' responses to one of the two preference options, by simply tossing a coin (40). Any of these manipulations of the data might be relatively harmless in terms of the final conclusions drawn from the testing if there were only a negligible amount of 'No Preference' responses. Yet, this is not always the case (36, 37, 41). Odesky (39) did provide some evidence that consumers who had been given a 'No Preference' option would, when denied this option, distribute their responses over the preference options in the same ratio as consumers who actually had had preferences. This surprising result was re-examined using the same stimuli as Odesky, as well as many other stimuli and using a larger sample of consumers (37). It was found not to be confirmed. Also, those researchers who eliminate the 'No Preference' responses or distribute them among the preferences, are indicating to the consumer that her opinion is not important and that she must say that she prefers a product chosen for her by the experimenter.

Besides providing more information (42), a further argument for the 'No Preference' option is that without it, it can be difficult to interpret the data. Should half the consumers report a preference for product 'A' and half report a preference for product 'B, no clear conclusion can be drawn if the 'No Preference' option were not available. It may be that all the consumers had no preference and chose the products randomly. Alternatively, it is also possible that half the consumers actually did prefer product 'A', while the other half preferred the product 'B'. The marketing decision taken for these two situations would be very different in each case. The 'No Preference' option attempts to avoid this problem. Thus, the arguments favor the use of a 'No Preference' option and abandonment of hopes of routine simple binomial analyses.

Besides providing a 'No Preference' option, other methods have been tried for detremining whether consumers have a preference or not. It can be argued that if a consumer changes her preference on repeated testing, she can be considered as not having an operational preference. This technique has been examined for preference tests with and without a 'No Preference' option. Chapman and Lawless (43) and Greenberg and Collins (44) used two tests in succession, Wilke *et al.* (45) used four while Baker *et al.* (46) used ten. All found high rates of inconsistency in the consumers' choices, indicating alternating preference choices from which a 'No Preference' judgment was deduced.

Statistical analysis For preference tests, one part of the statistical analysis is to determine an overall measure of the 'strength' of the preference in the sample of consumers. How far do they lean towards one product or the other? Another part is to determine whether the preference responses are different from a 'chance' situation where the consumers do not have a preference. Are the consumers responding in a way not determined by their preferences for the products or are they responding randomly or in a manner determined by the circumstances of the testing situation?

Analogies can be drawn with difference testing. To determine if the data indicate whether a consumer is performing better than mere guessing, a simple binomial statistical analysis can be used. To determine the extent of the perceived difference, Signal Detection Theory/Thurstonian modeling (1, 24, 47-49) provides a fundamental measure of degree of difference: d'. For preference testing there are two appropriate measures of degree of preference; the frequencies of response for each response option (Prefer A, No Preference, Prefer B) and the same fundamental measure used for difference testing: d'. In this case it denotes a degree of overall preference rather than a degree of difference. To establish whether the data differ from a situation where judges had no preference, a test of significance is required. As mentioned above, a binomial analysis, used for difference tests would be simple, hence the attempts described above to eliminate 'No Preference' responses. Yet, the data are not binomial; for 3 response categories (Prefer A, No Preference, Prefer B) the data are multinomial. Chi-squared is a simple multinomial test and usage of this analysis has been investigated (36, 38). For such an analysis, the response frequencies for the two different products to be compared for preference, provide the 'observed' frequencies. However, there is a problem with the 'expected' frequencies

For a chi-squared test, expected frequencies are usually generated using a null hypothesis. Yet, to test observed frequencies against expected frequencies generated on a null hypothesis, would be to test whether the data differ from an equal distribution of cases in each of the 3 categories: 'Prefer A', 'Prefer B', 'No Preference'. Such an equal distribution would represent a situation where a third of the consumers had no preference, a third preferred 'A' and a third preferred 'B'. It could also represent a situation where consumers ignored the products, closed their eyes and randomly marked one of the three categories on their response sheets. Neither of these possibilities corresponds to the case where consumers tasted the products and on reflection decided that they had no preference.

Logically, the expected frequencies for the case where there is not a preference, would be 100% in the 'No Preference' category and zero in both preference categories. Yet, the testing situation induces false preference responses from consumers who have, in fact, no preference. This was first reported by Ennis in 1980 (50) when he noted that consumers expressed preferences when presented with identical pairs of cigarettes. His samples of consumers made various analytical judgments as well as being asked for their preferences. He noted that the best overall fit to the data was that only 20% of consumers reported having no preference, while the remaining 80% distributed their responses equally between the two products, producing a 40-20-40 split. He reports having obtained the same results in subsequent unpublished testing.

This initiated research by a variety of researchers who published response frequencies for selections of putatively identical stimuli. They varied with the products being tested, how the response options were phrased and the number available, the experimental conditions and the origin of the consumers; tests have been conducted in USA, Mexico and Korea (36-38, 43, 51). Foley *et al.* (52) also made such a study but their response frequencies are not published. In such studies, the 40-20-40 ratio was rarely found. It is interesting that Korean consumers appeared more prone to the bias of having preferences for identical stimuli than Americans (38).

The placebo pair The data from putatively identical stimuli can be envisaged like a placebo in medical research. It provides the response frequencies that are not the result of preferences of the products being tested; they are the preference responses elicited by the conditions of the testing situation. For data analysis they have been used in two ways (36, 53). Firstly, they have been used as the expected frequencies in a chi-squared analysis. These can be derived either from a separate control sample of consumers (38, 52) or from the same sample of judges being used to assess the different products under consideration (36, 53). The second strategy would seem a better choice because it avoids the assumption that the samples of consumers testing the 'identical' samples and those testing the different products are equivalent. Yet, the approach poses problems of whether the 'identical' samples should be tasted before or after the different samples under consideration. Alfaro-Rodriguez *et al.* (36) found that it made little difference but Kim *et al.* (53) did find some context effects for products that were visually easy to discriminate. This topic is currently being researched.

There is a second use for the placebo data; it can be used as a screening tool to select only those consumers who, at the time of testing, had not been affected by the hidden demand characteristics of the testing situation. They would be responding only in terms of their preference (36, 53). The chi-squared analysis does not eliminate the bias elicited by the testing conditions; it merely indicates whether the responses are different from a 'No Preference' situation. The bias still remains when the different products under consideration are being tested. Some of the preferences will still be false. If the placebo 'No Preference' responses are used as a screening tool to select those consumers who are unbiased by the testing conditions, then their preference responses could be considered to be true preferences. Yet, this latter approach has the disadvantage of greatly reduced sample sizes. Kim *et al.* (53) reported a yield of only 30% of consumers while Alfaro-Rodriguez *et al.* (36) reported even lower numbers. It is therefore useful to investigate ways of increasing the proportion of 'No Preference' responses in the placebo condition, to attempt to increase the proportion of consumers who would pass the screening test.

Marchisano *et al.* (38) and Kim *et al.* (53) had experimented with a variety of response options for 'identical' stimuli in the placebo paired preference test. Generally, they found that increasing the number and variety of 'No Preference' options increased the frequency of 'No Preference' responses. Chapman *et al.* (51) also found a similar result. Yet, a consideration of whether preference questions are really the correct questions to ask might lead to some solutions. Preference has usually been measured in terms of liking more. Yet, the question more relevant to marketing is whether the consumer is more likely to buy the product. Kim *et al.* (53) introduced elements of buying into the preference response options and found a slight increase in 'No Preference' responses. Yet, current research is examining more extreme options like "I would only buy 'A' and never buy 'B' vs. "depending on price, availability and my mood etc., I might buy either product", the latter might well be chosen with a frequency well above 30%. This and other milder response options are at present being investigated.

Summary

From the above discussion, it can be seen that even a simple thing like a paired preference test which must give accurate data for marketing decisions, becomes complicated when it is examined closely. This is true of nearly all measurement techniques used in consumer and marketing research. Too many assumptions have been made about measurement techniques and these assumptions are being tested now and found wanting. This is why consumer/ sensory science is necessary for marketing. It ensures that behavioral measurement methods are examined and tested for their validity. The techniques and knowledge of the consumer/sensory scientist can be a valuable asset to provide guidance regarding the collection of accurate data essential for successful marketing.

References

1. Lee H-S, O'Mahony M. Sensory difference testing: Thurstonian models. Food Sci. Biotechnol. 13: 841-847 (2004)
2. Lee H-S, O'Mahony M. The evolution of a model: A review of Thurstonian and conditional stimulus effects on difference testing. Food Qual. Prefer. 18: 369-383 (2007)
3. MacFie H. Preference mapping and food product development. pp. 551-592. In: Consumer-led Food Product Development. MacFie H (ed). CRC Press, New York, NY, USA (2007)

4. Delwiche JF. Thurstonian probabilistic approaches to new food product development. pp. 456-470. In: Consumer-led Food Product Development. MacFie H (ed). CRC Press, New York, NY, USA (2007)
5. Kim NS. Discriminant analysis of marketed liquor by a multi-channel taste evaluation system. Food Sci. Biotechnol. 14: 554-557 (2005)
6. Chung SJ. Flavor release from ice cream during eating. Food Sci. Biotechnol. 16: 8-17 (2007)
7. Lee O-H, Lee H-S, Lee S-M, Kim Y-K, Kim K-O. Sensory characteristics and consumer acceptability of various green teas. Food Sci. Biotechnol. 17: in press (2008)
8. Rosas-Nexticapa M, Angulo O, O'Mahony M. How well does the 9-point scale predict purchase frequency? J. Sens. Stud. 20: 313-331 (2005)
9. Peryam DR, Girardot NF. Advanced taste-test method. Food Eng. 24: 58-61, 194 (1952)
10. Peryam DR, Pilgrim FJ. Hedonic scale method of measuring food preferences. Food Technol.-Chicago 11: 9-14 (1957)
11. Peryam DR, Polemis BW, Kamen JM, Eindhoven G, Pilgrim FJ. Food preferences of men in the U.S. Armed forces. Department of the Army Quatermaster Research and Engineering Command-Quatermaster Food and Container Institute for the Armed Forces, Chicago, IL, USA (1960)
12. Jones LV, Thurstone LL. The psychophysics of semantics: An experimental investigation. J. Appl. Psychol. 39: 31-36 (1955)
13. Jones LV, Peryam DR, Thurstone LL. Development of a scale for measuring soldiers' food preferences. Food Res. 20: 512-520 (1955)
14. Edwards AL. The scaling of stimuli by the method of successive intervals. J. Appl. Psychol. 36: 118-122 (1952)
15. Jeon SY, O'Mahony M, Kim KO. A comparison of category and line scales under various experimental protocols. J. Sens. Stud. 19: 49-66 (2004)
16. Kim KO, O'Mahony M. A new approach to category scales of intensity I: Traditional versus rank rating. J. Sens. Stud. 13: 241-249 (1998)
17. Koo TY, Kim KO, O'Mahony M. Effects of forgetting on performance on various intensity scaling protocols: Magnitude estimation and labeled magnitude scale (Green scale). J. Sens. Stud. 17: 177-192 (2002)
18. Lee HJ, Kim KO, O'Mahony M. Effects of forgetting on various protocols for category and line scales of intensity. J. Sens. Stud. 16: 327-342 (2001)
19. Lawless HT. Contextual effects in category ratings. J. Test. Eval. 11: 346-349 (1983)
20. Lawless HT, Malone GJ. A comparison of rating scales: Sensitivity, replicates, and relative measurement. J. Sens. Stud. 1: 155-174 (1986)
21. Parducci A. Range-frequency compromise in judgement. Psychol. Monogr. 77: 1-50 (1963)
22. Parducci A. Category judgement-A range-frequency model. Psychol. Rev. 72: 407-418 (1965)
23. Pipatsattayanuwong S, Lee HS, Lau S, O'Mahony M. Hedonic R-index measurement of temperature preferences for drinking black coffee. J. Sens. Stud. 16: 517-536 (2001)
24. Green DM, Swets JA. Signal Detection Theory and Psychophysics. John Wiley and Sons, New York, NY, USA. pp. 45-50 (1966)
25. Brown J. Recognition assessed by rating and ranking. Brit. J. Psychol. 65: 13-22 (1974)
26. Lee HS, O'Mahony M. Sensory evaluation and marketing: Measurement of a consumer concept. Food Qual. Prefer. 16: 227-235 (2005)
27. Karahadian C. Impact of global markets on sensory testing programs. Food Technol.-Chicago 49: 77-78 (1995)
28. Stone H, Sidel JL. Strategic applications for sensory evaluation in a global market. Food Technol.-Chicago 49: 80-89 (1995)
29. Sobal J. Cultural comparison research designs in food, eating, and nutrition. Food Qual. Prefer. 9: 385-392 (1998)
30. Yeh LL, Kim KO, Chompreeda P, Rimkeeree H, Yau NJN, Lundahl DS. Comparison in the use of the 9-point hedonic scale between Americans, Chinese, Koreans, and Thai. Food Qual. Prefer. 9: 413-419 (1998)
31. Prescott J, Laing D, Bell G, Yoshida M, Gillmore R, Allen S, Yamazaki K, Ishii R. Hedonic responses to taste solutions: A cross-cultural study of Japanese and Australians. Chem. Senses 17: 801-809 (1992)
32. Yao E, Lim J, Tamaki K, Ishii R, Kim KO, O'Mahony M. Structured and unstructured 9-point hedonic scales: A cross-cultural study with American, Japanese, and Korean consumers. J. Sens. Stud. 18: 115-139 (2003)
33. Stone H, Sidel JL. Sensory Evaluation Practices. Academic Press, San Francisco, CA, USA. pp. 246-251 (1992)
34. Lawless HT, Heymann H. Sensory Evaluation of Food: Principles and Practices. Chapman and Hall, New York, NY, USA. pp. 431-449 (1998)
35. Resurreccion AVA. Consumer Sensory Testing for Product Development. Aspen Publ., Gaithersburg, MD, USA. pp. 10-20 (1998)
36. Alfaro-Rodriguez H, Angulo O, O'Mahony M. Be your own placebo: A double paired preference test approach for establishing expected frequencies. Food Qual. Prefer. 18: 353-361 (2007)
37. Angulo O, O'Mahony M. The paired preference test and the 'No Preference' option; was Odesky correct? Food Qual. Prefer. 16: 425-434 (2005)
38. Marchisano C, Lim J, Cho HS, Suh DS, Jeon SY, Kim KO, O'Mahony M. Consumers report preference when they should not: A cross-cultural study. J. Sens. Stud. 18: 487-516 (2003)
39. Odesky SH. Handling the neutral vote in paired comparison product testing. J. Marketing Res. 4: 199-201 (1967)
40. Alfaro-Rodriguez H, O'Mahony M, Angulo O. Paired preference tests: d' Values from Mexican consumers with various response options. J. Sens. Stud. 20: 275-281 (2005)
41. Falk SN, Henrickson RL, Morrison RD. Effect of boning beef carcasses prior to chilling on meat tenderness. J. Food Sci. 40: 1075-1079 (1975)
42. Gridgeman NT. Pair comparison, with and without ties. Biometrics 15: 382-388 (1959)
43. Chapman KW, Lawless HT. Sources of error and the no-preference option in dairy product testing. J. Sens. Stud. 20: 454-468 (2005)
44. Greenberg A, Collins S. Paired comparison taste tests: Some food for thought. J. Marketing Res. 3: 76-80 (1966)
45. Wilke KD, Cochrane C-YC, Chambers IV E. Multiple preference tests can provide more information on consumer preferences. J. Sens. Stud. 21: 612-625 (2006)
46. Baker GA, Amerine MA, Roessler EB, Filipello F. The nonspecificity of differences in taste testing for preference. Food Res. 25: 810-816 (1960)
47. Macmillan NA, Creelman CD. Detection Theory: A User's Guide. 2nd ed. Lawrence Erlbaum Associates, Mahwah, NJ, USA. pp. 7-16 (2005)
48. Thurstone LL. A law of comparative judgment. Psychol. Rev. 34: 273-286 (1927)
49. Thurstone LL. Psychophysical analysis. Am. J. Psychol. 38: 368-389 (1927)
50. Ennis DM. A common framework for interpreting similarity, liking, and preference data (abstract no. CS3.2). In: Abstracts: 7th Pangborn Sensory Science Symposium. August 12-16, Hyatt Regency, Minneapolis, MN, USA. Elsevier, Oxford, UK (2007)
51. Chapman KW, Grace-Martin K, Lawless HT. Expectations and stability of preference choice. J. Sens. Stud. 21: 441-455 (2006)
52. Foley M, Williams A, Bade J, Lancaster B, Popper R, Carr BT. Effect of preference-question format with and without sample differences (abstract no. 043). In: Abstracts: 5th Pangborn Sensory Science Symposium. July 20-24, The Boston Park Plaza Hotel, Boston, MA, USA. Elsevier, Oxford, UK (2003)
53. Kim HS, Lee HS, O'Mahony M, Kim KO. 'Preference' and 'No Preference' responses when identical pair and non-identical pair was presented in different test protocols. J. Sens. Stud. 22: in press (2007)