Available online at www.sciencedirect.com
ScienceDirect
Food Quality and Preference 18 (2007) 89–96
Food Quality and Preference
www.elsevier.com/locate/foodqual

Relating consumer and trained panels' discriminative sensitivities using vanilla flavored ice cream as a medium

R. Ishii [a], H. Kawaguchi [a], M. O'Mahony [a], B. Rousseau [b,*]

[a] *Department of Food Science and Technology, University of California, Davis, CA 95616, United States*
[b] *The Institute for Perception, 2306 Anza Avenue, Davis, CA 95616, United States*

Received 19 June 2004; received in revised form 24 July 2005; accepted 22 August 2005
Available online 15 December 2005

Abstract

This research investigated the possibility of uncovering a relationship between the sensitivities of trained and consumer panels in three experiments. In Experiment I, five studies were conducted using vanilla flavored ice cream. In each study, two ice cream samples differing in formulation and/or their manufacturing process were used. They were compared by both panels using same–different tests with sureness judgments (degree of difference methodology). Using the appropriate probabilistic Thurstonian model, d' values, a measure of the underlying sensory difference perceived between the products, were calculated and the underlying relationship between the two panels' sensitivities uncovered. An additional study was then conducted (Experiment II). A new pair of ice creams differing in fat content was first evaluated by the trained panel. Based on the estimated d' value (trained panel's measured $d' = 2.69$), the corresponding consumer d' value was predicted using the relationship uncovered in Part I (consumer panel's predicted $d' = 1.54$). The same pair of ice creams was then evaluated by the consumer panel. The measured d' value for the consumers was almost identical to that predicted by the uncovered relationship (consumer panel's measured $d' = 1.56$). Along with the discrimination component of these studies, paired preference tests were conducted (Experiment III) in order to study a link between perceived difference and expressed preference. The results give an indication of when a perceived difference might start translating into a change in acceptability of the original product. These results indicate the potential of such an approach to predict consumers' perceptions from an in-house semi-trained or trained panel, providing a useful predictive tool and a means of reducing repetitive and costly consumer testing.

Keywords: Consumers; Experts; Thurstonian; Sensitivity; Relationship

1. Introduction

One objective of sensory evaluation is to obtain information regarding how consumers perceive products. A second objective is to measure consumers' likes, dislikes and preferences. In both cases consumers are sampled according to the appropriate demographic criteria and given the appropriate tests. They are not given any special training with the food, so as to avoid them becoming more sensitive to changes in the food attributes than would be appropriate to consumers. The first objective has been described as Sensory Evaluation II and the second as Consumer Acceptance and Preference testing (O'Mahony, 1995). Both are expensive and time consuming, requiring recruitment, screening and testing of consumers under conditions that come as close to realistic as possible.

On the other hand, more analytical work is conducted 'in-house' with judges who have some degree of training. This has been called Sensory Evaluation I (O'Mahony, 1995). Subjects are often company employees or individuals who have been specially screened, recruited and trained for such studies. For descriptive analysis, the use of trained panelists is recommended because naïve consumers usually lack the experience, vocabulary and concept alignment necessary to generate quality descriptive data. A few studies have investigated the possibility of using consumers in

* Corresponding author. Tel.: +1 530 753 2231; fax: +1 530 757 1010.
E-mail address: brousseau@sbcglobal.net (B. Rousseau).

0950-3293/$ - see front matter © 2005 Elsevier Ltd. All rights reserved.
doi:10.1016/j.foodqual.2005.08.004

descriptive analysis, but the ability of such untrained panelists has never been successfully demonstrated (Cardello et al., 1982; Chollet & Valentin, 2001; Dugle, 1997; Gonzàlez, Benedito, Carcel, & Mulet, 2001; Hough, 1998; Lesschaeve & Issanchou, 1995; Moskowitz, 1996, 1997, 1998; Muñoz & Chambers, 1993; Muñoz, Chambers, & Hummer, 1996). Even QDA, which is conceived as representing consumer perceptions, requires experienced consumers to go through a concept alignment exercise (Stone & Sidel, 1993).

It is generally assumed and accepted that consumers are less sensitive or discriminating than a panel of appropriately trained subjects. Research has been conducted comparing relative sensitivities and discrimination abilities of trained and naïve panelists. While no differences in olfactory detection threshold were identified between the two types of subjects (Bende & Nordin, 1997; Parr, Heatherbell, & White, 2002), Solomon (1990) showed that trained subjects elicited a higher performance than novices. One explanation for these results is that trained panelists are more familiar with the experimental procedures, which in turn allows them to discriminate better among the stimuli under study. However, Chollet, Valentin, and Abdi (1997) found that such superior performance was limited to familiar stimuli.

These results can be adjusted based on the consumers' consumption status, as individuals who are heavy users of a given product could potentially discriminate as well as, or even better than, a panel of trained subjects. Most of the experimental investigations cited earlier focused on the ability of consumers to rate sensory attributes or articulate their perceived sensations. None of them studied consumers' and trained panels' sensitivities as measured using discrimination methodology. Such knowledge would be particularly valuable considering the way the results of discrimination methods are frequently utilized in the industry. For example, when a trained panel detects a difference between a standard product and its reformulation, the approach is usually conservative and the reformulation is not released in the market, even though it could have been a source of significant cost savings. The logic here is that consumers *might* detect the formulation change. On the other hand, they might not and a significant cost reduction opportunity may have been lost. Also, a trained panel might miss differences that heavy users of the product could actually pick up. Based on these ideas, it seems that the knowledge of the relationship between the sensitivity of a company's trained panel and that of the consumer group most important to a company's business, would be especially valuable.

As mentioned above, the results available in the literature focused mainly on the rating of attributes by consumer and trained panels. Some of the approaches involved the estimation of the spread of the products' ratings (Moskowitz, 1996), average random error variance (Gonzàlez et al., 2001), ability to discriminate using descriptive attributes vs. hedonic attributes (Chollet & Valentin, 2001) and aptitude to describe product characteristics (Lesschaeve & Issanchou, 1995). Moskowitz (1996) also proposed a technique that would allow the prediction of consumers' results from trained panelists data and vice-versa. The validity of his approach was later fervently challenged (Dugle, 1997; Hough, 1998). However, none of these approaches were based on a model of actual sensory perception.

In order to build a valuable tool that would allow a sensory scientist to relate the sensitivities of consumer and trained panels, a theory of product perception is necessary. Available models are those based on the pioneering work of Thurstone (1927). A key aspect of Thurstonian modeling is that, everything else being equal, the measure of the degree of difference between products is independent of the methodology used (see for instance Braun, Rogeaux, Schneid, O'Mahony, & Rousseau, 2004; Ennis, 1993; Frijters, 1980; Hautus & Irwin, 1995; O'Mahony, Masuoka, & Ishii, 1994; Rousseau & O'Mahony, 1997; Stillman, 1993). Thurstonian modeling is based on two ideas: (1) Stimulus perception varies from evaluation to evaluation due to nerve impulse variations or variability in the stimulus itself; (2) Subjects use specific decision rules to perform the task they were asked to execute (difference test, rating, etc.) With such assumptions, an index of sensory difference called d' can be estimated; it corresponds to the distance between the means of the perceptual distributions measured in terms of their standard deviation. A value of unity is approximately equivalent to threshold: 75% of 2-AFC tests correct. This concept is illustrated in the upper part of Fig. 1.

The effect of practice can be modeled using the Thurstonian approach. As mentioned above, a d' value is measured in terms of the distributions' standard deviation, i.e. variance. For two given products A and B with perceptual means μ_A and μ_B evaluated in two different testing conditions, different d' values will be obtained if the variance of the perceptual distributions in each condition are different. Using this model, practice can be modeled by a reduction of the perceptual variance. As the size of the measuring unit decreases, a larger d' value will be measured

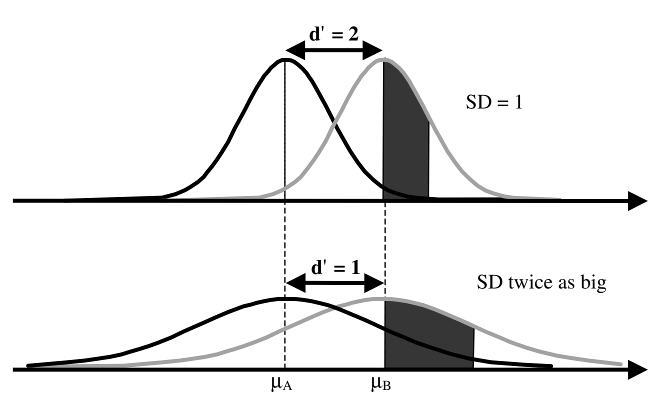

Fig. 1. Thurstonian representation of the effect of training on perceptual sensitivity.

using a group of trained panelists rather than a group of naïve consumers, if these consumers are indeed less discriminating than the trained panelists. This concept is shown in Fig. 1. Such increase of performance by reducing the distributions' perceptual variance has been modeled and observed experimentally for the issue of stimulus resampling (Juslin & Olsson, 1997; Rousseau & O'Mahony, 2000).

It is worth noticing that such an increase in performance through practice is possible as long as there is potential for the perceptual variance to decrease through factors such as attention, concentration and familiarity with the testing procedure. Studies involving texture have shown for instance that trained panelists have more constant chewing patterns than untrained subjects across samples and sessions (Gonzàlez, Sifre, Benedito, & Nogués, 2002; Mioche & Martin, 1998). However, if the main source of variability comes from the stimulus itself, e.g. highly variable products such as fruits, potato chips, etc., practice might not be successful at reducing the perceptual variance, resulting in a trained panelist's sensitivity similar to that of untrained consumers.

The project described here investigated the possibility of relating the sensitivity of a trained panel with that of a group of consumers using vanilla flavored ice cream samples. It involved three experiments. In Experiment I, five studies were conducted to generate the relationship between the two panels. In Experiment II, an additional study was carried out to test the predictability of the uncovered relationship. In Experiment III, additional preference data was collected in order to explore the degree of difference above which a preference might be expressed.

2. Experiment I

Five studies involving pairs of ice-cream samples of various degrees of similarity were conducted. Each pair of ice-creams was evaluated by both a consumer and a trained panel. This first experiment was conducted in order to establish the potential relationship between the two panels' sensitivities.

2.1. Materials and methods

2.1.1. Subjects

2.1.1.1. Trained panel. Subjects were students, staff and their friends from the University of California at Davis. They were first recruited based on their regular consumption of ice-cream products, their interest in being part of an ice-cream discrimination panel and their availability over a sufficiently extended period of time. A pool of approximately 20 trained panelists was available at any time during the project. Twelve to eighteen panelists from this pool of judges were used in each of the 5 individual studies.

After they had been recruited, they followed a screening and training procedure to establish that they possessed the appropriate discrimination ability to be part of the panel. The screening and training process involved the use of several pairs of ice creams samples of different similarity levels that the panelists were expected to be able to discriminate. Some of the differences involved variations in fat content, which created differences both in texture and flavor characteristics. The methodology used was the same as that used in this project, namely the same-different procedure with sureness judgments. Prospective panelists not showing sufficient performance improvement as measure by his/her ability to discriminate between two particular ice-cream samples were not retained to be part of the final panel pool. This screening/training lasted between one week and one month depending on the panelist, with up to four sessions of one hour per week. The total training time varied between 4 and 16 h depending on the subject's sensitivity.

Panelists who were recruited, participated in regularly scheduled sessions with a minimum of one per month for this and other relevant projects. Some of the projects involved the use of discrimination methodologies rather than traditional rating methods as a means of conducting descriptive analysis projects. If no study was scheduled in a particular month, a training session was organized to ensure that the panelists retained their discriminative ability.

2.1.1.2. Consumer panel. Subjects were recruited from the campus of the University of California at Davis. They were students, staff, faculty, friends and visitors intercepted by the Department of Food Science and Technology. A different pool of subjects was used in each of the five studies.

Table 1 summarizes the demographics of the panels involved in the five studies.

2.1.2. Stimuli

Stimuli were vanilla flavored ice-creams provided by a major ice cream manufacturer. Differences between the samples to be compared varied from texture to flavor, following process or formulation alterations. They were chosen so as to represent sensory differences of different magnitudes. The samples of ice-cream were scooped by a single person to ensure consistency of serving. The top, peripheral and bottom parts of the ice-cream container were not sampled in the experiment to ensure consistency of ice cream from container to container. Samples were scooped and placed in 6 oz Styrofoam cups and kept on trays in a horizontal freezer at approximately 0–5 °F. Cups

Table 1
Trained panelist and consumer demographics in studies 1–5

	Trained panelists	Consumers
Study #1	18 (7 M, 11 F; 20–68 years)	133 (55 M, 78 F; 16–60 years)
Study #2	17 (5 M, 12 F; 20–49 years)	124 (55 M, 69 F; 17–66 years)
Study #3	16 (4 M, 12 F; 20–49 years)	122 (49 M, 73 F; 17–66 years)
Study #4	12 (7 M, 5 F; 23–68 years)	137 (59 M, 78 F; 17–66 years)
Study #5	13 (6 M, 7 F; 23–40 years)	232 (99 M, 133 F; 17–60 years)

were taken out of the freezer immediately prior to evaluation by the panelists. Each portion comprised approximately 21 g of ice-cream.

2.1.3. Procedure

2.1.3.1. Trained panel. Panelists were placed in isolated booths, with a maximum of four panelists present simultaneously. Upon arrival in the booth, a panelist would first rinse his/her mouth with cold tap water. A "warm-up" pair of the two ice-creams under study (A and B) was then presented in order for the panelist to become familiar with the difference to be sought. Then two rinses with warm tap water (approximately 100 °F) were taken to eliminate the fattiness sensation of the samples tasted, followed by a rinse of cold tap water to reduce the mouth temperature to a more compatible level. The protocol used here was the same–different method with sureness judgments (sometime called 'Degree of difference methodology'). Eight successive pairs of ice-creams were evaluated, four being made of identical ice-creams (two AA and two BB), the other four of different ice-creams (two AB and two BA). The panelists were not aware of this arrangement. For both the warm-up and actual test pairs, the panelist could taste as little or as much of the ice creams as needed. The eight pairs were presented in a random order that varied from subject to subject. For each pair, the panelist was required to state whether the two ice-creams were the same or different, and whether he/she was sure or unsure of his/her answer. The same rinsing protocol as before was used in between successive pairs. Stimuli were presented in sets of two pairs on a tray (total 4 trays).

Samples were evaluated under red light to mask any potential visual difference between the samples. Sessions lasted between 10 and 30 min depending on the panelist and product comparison.

2.1.3.2. Consumer panel. Upon interception on the campus, a consumer would be escorted to the lobby of the Department of Food Science and Technology where the study was being conducted. Demographic information was collected and the subject was asked to rinse his/her palate with cold water provided in a cup.

Subjects were asked to perform a same-different test between the two samples, indicating if they thought the samples were the same or different, and whether they were sure or not sure of their answer (degree of difference methodology). They could taste the samples as many times as they wished, the only limitation being the amount of ice-cream provided (approximately 21 g). Responses and any comments were given verbally.

Very slight visual differences were sometime visible. Red light could not be used in this particular setting. Consequently, consumers were asked not to pay attention to any potential visual differences. In order to ensure that they would focus on aroma, flavor and texture attributes only, they were told that even two samples of ice-cream scooped from the same container could show visual differences. As the interview was conducted one to one, the interviewer could also observe whether the consumer was paying too much attention to appearance. Throughout the seven studies, this was never observed.

Testing including establishment of rapport, taking of demographic details etc., lasted 10–15 min depending on the consumer and pairs of products involved.

2.2. Results

For each study and panel, a d' value with its associated variance was calculated using the IFPrograms software (The Institute for Perception, Richmond, VA, USA.) The analysis involved treating the responses as difference ratings (see O'Mahony & Rousseau, 2003) and using the method of maximum likelihood to estimate d', τ (the cognitive criterion used by the subject to generate an answer (Ennis, Palen, & Mullen, 1988; Rousseau, 2001)) and the variance–covariance matrix for these estimates. Table 2 summarizes the results of the five studies. Significant differences indicate that the trained panel was significantly more discriminating than the consumer panel.

2.3. Studying the trained panelist/consumer relationship

A regression analysis was conducted between the two sets of d' values and the relationship is illustrated in Fig. 2. The high r^2 value indicates that the relationship

Table 2
Consumer and trained panelist d' values, with significance levels for the difference between the two d' values

	Consumers	Trained panelists	p Value
Study #1	0.00	1.10	0.01
Study #2	0.82	1.91	<0.01
Study #3	1.57	2.08	0.18
Study #4	2.63	4.30	<0.01
Study #5	2.70	4.10	0.02

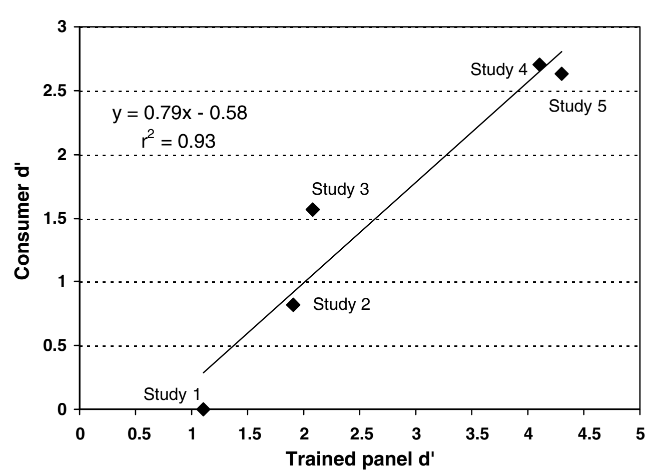

Fig. 2. Relationship between the trained and consumer panels' sensitivities based on five studies

should be quite successful at predicting consumers' discrimination ability from a trained panel's d' value. From the graph (Study 1) it can be seen that when the trained panelists were discriminating at around threshold levels ($d' = 1$), consumers could not discriminate ($d' = 0$). For threshold discrimination with consumers, trained panelists were discriminating well ($d' = 2$). A discrimination level of $d' = 2$ for the trained panel was used as a decision rule in subsequent projects not specifically related to this research. Therefore, if a d' of 2 or higher was estimated, the recommendation would be to not go forward with the formulation or process change as consumers might be able to discriminate between the samples. A change inducing a d' of 2 or lower would be interpreted as acceptable as the potential discrimination of the samples by consumers would be limited, if at all possible.

3. Experiment II

In this experiment, an additional study involving trained and consumer panels was conducted in order to corroborate the value of the relationship uncovered in Experiment I.

3.1. Materials and methods

3.1.1. Subjects

3.1.1.1. Trained panel. The trained panel comprised 17 subjects (5 M, 12 F, 20–49 years) sampled from the same group of trained panelists as those used in studies 1–5.

3.1.1.2. Consumer panel. One hundred twenty (120) subjects (63 M, 57 F, 17–66 years) were again recruited from the campus of the University of California at Davis. They were students, staff, faculty, friends and visitors intercepted by the Department of Food Science and Technology.

3.1.2. Stimuli

Stimuli were two vanilla flavored ice creams differing in fat content and provided by the same major ice cream manufacturer.

3.1.3. Procedure

The procedures followed were the same as those used in the first five studies for both the trained and consumer panels.

3.2. Results

The d' values were again estimated using the IFPrograms software (The Institute for Perception, Richmond, VA, USA). The trained panel's d' was found to be 2.69 (variance = 0.09). Using the relationship uncovered in Fig. 2, the consumer would be predicted to exhibit a d' value of 1.54 (95% confidence limits of [0.28; 2.79]). Upon analyzing the consumers' data, a d' of 1.56 (variance = 0.05) was estimated. Again, the trained panelists were found to be significantly more sensitive than the consumers ($p = 0.003$).

4. Experiment III

Along with a difference test, consumers in studies 2–6 performed a preference test between the pair of ice creams being investigated (as the earliest study, Study 1 did not have this aspect included in its design.) This information was collected in order to investigate how large of a sensory difference will alter consumers' preferences. In turn, this provides further insights to establish the d' value above which a sensory difference really matters to the consumer.

4.1. Materials and methods

4.1.1. Subjects

The consumers' demographics were the same as those described in studies 2–6 in Experiments I and II.

4.1.2. Stimuli

Stimuli were the same as those used in studies 2–6 in Experiments I and II.

4.2. Procedure

Prior to performing the same-different pairs in each of the studies, consumers were presented with the two ice-creams and asked whether they preferred one or the other, or whether they did not have a preference.

4.3. Results

Results are summarized in Table 3. Because the interest here was only in those consumers who had expressed a preference, the numbers who had expressed 'No preference' was noted but not included in the analysis. Using just those

Table 3
Preference test results for five of the six consumer studies

	Preference A	No preference	Preference B	p	% Expressed preference	Trained panel d' from difference study
Study #2	44	29	51	NS	77	1.91
Study #3	63	19	40	0.03	84	2.08
Study #4	40	12	81	<0.001	91	4.30
Study #5	88	26	118	0.04	89	4.10
Study #6	47	29	44	NS	76	2.69

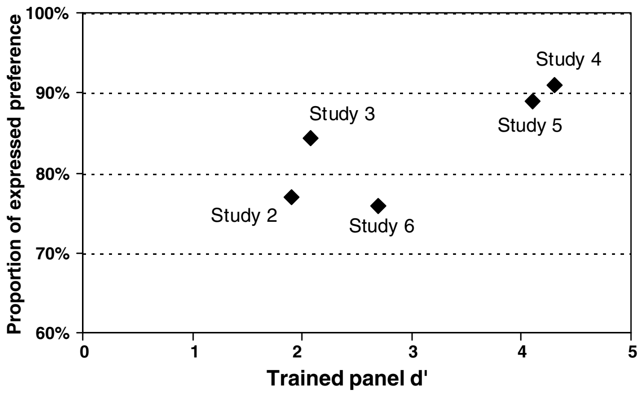

Fig. 3. Expert panel d' values vs. consumers' proportions of expressed preference between the products.

consumers who had expressed a preference a binomial analysis was performed.

The relationship between the trained panel d' value and the percentage of consumers expressing a preference is shown in Fig. 3.

5. Discussion

The procedure described here indicates that it is possible to study and find a relationship between the sensitivities of trained and consumers panels. Once this relationship is uncovered, it can be used to make predictions on how large or small of a difference consumers will perceive based on results obtained from a trained panel. The first thing that can be noted is that the trained panel exhibited a significantly higher sensitivity than the consumers in five of the six studies, the last study (Study #3) indicating the same—non-significant—trend. This result indicates that proper training and testing conditions can significantly improve subjects' performance compared to that of naïve consumers in the investigation of commercial ice-cream properties.

The prediction found here is surprisingly accurate and such accuracy should probably not be expected every time. Any prediction has variability associated to it and the predicted value of 1.54 has a 95% confidence interval of [0.28; 2.79]. Involving a greater number of studies would permit to predict consumer discrimination abilities with greater accuracy. For instance, including the results of Study 6 in the regression analysis would tighten the 95% confidence limits to [0.61; 2.48] for the same 1.54 prediction. Nevertheless, the current results underscore the value and usefulness of such an approach to investigate and predict consumers' sensitivity.

This approach is not limited to the use of the degree of difference method. One of the benefits of the Thurstonian approach is the possibility of using almost any type of protocol (as long as the responses are categorical) to generate d' values. Numerous discrimination methods could be utilized, such as the triangle, duo-trio, 2-AFC, 3-AFC, tetrads tests, etc, as well as rating methods with any number of response categories (5, 7, 9, etc.) Therefore, a relationship could be studied using the same methodology twice, as was done in the present study, or with a mixture of methods, such as a same–different paradigm with consumers and a triangle method with a trained panel.

Thurstonian modeling can explain reduced discrimination (lower d') in terms of an increase in perceptual variance (cf. Fig. 1). This theory provides a simple model for prediction, but it is worth noting that its assumptions have not been specifically corroborated in this research. To do so, replicated measurements on all—consumer and trained panelist—individuals would be necessary to confirm that the variability associated with the trained panel's data is smaller than that of the consumer panels. Nevertheless, differences in perceptual variability provide a potential and sensible explanation to the results observed here. Further research will be necessary to confirm the Thurstonian explanation.

Furthermore, it could be argued that if variance were the only variable, then when consumers had a d' of zero, trained panelists should also show no discrimination. Consequently, the relationship between the two panels should go through the origin of the graph. This is not the case here (cf. Fig. 2). Some support for this position could be found by considering that, as the sensory difference gets smaller, variance is added due to a greater variability in the τ-criterion for consumers, the psychological parameter inherent to the same-different and degree of difference methodologies (Ennis et al., 1988; Rousseau, 2001). As trained panelists are more accustomed to the procedure, and their perceived difference (d') is larger, such variation in τ criterion would be more limited.

Another candidate is attention. It can be argued that consumers do not always attend to the relevant attribute in discrimination tests. Slight differences in ice creams that varied in texture would not be detected by consumers if they only attended to flavor. In Thurstonian terms this can be seen as consumers not selecting the appropriate dimension along which the stimuli were varying. The smaller the difference between the products, the more likely this situation is to arise.

The uncovered link between the two types of panel is of course closely related to the specifications of this project. Alternative relationships could and would be found depending on the amount of training the trained panel is exposed to, or if a different group of consumers such as heavy users of a given product were to be used. As mentioned earlier, the amount of variation associated with the product itself might prevent any improvement through training, if it happens to be the main source of perceptual variance.

It is worth considering the level of discrimination (d' value) above which a project manager would consider that the sensory difference between the products is no longer negligible for the consumer. For the present study, while the trained panelists exhibit a discrimination level corresponding to a d' of 1, the consumer group would not be seen as indicating perception of a difference between the products. Therefore, if a difference with a d' less than 1 is found with the trained panel, the project manager could conclude that this difference would be negligible to the consumers. This might not have been his/her conclusion had the trained panel's result been the only source of information available.

Another approach to the perception of a reformulation or ingredient change in a product would involve the collection of additional information regarding consumers' preferences. Consumers may perceive a slight change but might not mind it. It can be argued that a sensory difference becomes important if it results in a development of a preference for one or the other of the two products. Along with a difference test, the consumer could be presented with a pair of the two samples and asked whether he/she has a preference between the two, as was done in Experiment III. The discrimination d' value above which a significant preference develops can be then used as the representative threshold. Results from Experiment III of this project provided some insight on the feasibility of this approach (cf. Table 3 and Fig. 3).

The relationship illustrated in Fig. 3 indicates, as expected, that as the size of the sensory difference increases, the percentage of expressed preference increases. It can then be used to set an upper limit on the size of a meaningful sensory difference based on preference data. It is worth noticing that it appears that for a d' of 0 (no sensory difference), one would probably not observe 0% of expressed preference. This confirms other research that showed that the proportion of 'no preference' answers when presenting identical products will range from 8% to 30% (Ennis, 2001; Marchisano et al., 2003).

One last piece of information that could be gathered is consumer complaints over the years following changes made to the product formulation or manufacturing process. However, these types of data might be more difficult to obtain, and less reliable, than the methods described earlier.

The approach outlined in this study can significantly alter the way the sensory scientist and his/her management make decisions on the appropriateness of a process or formulation change in a company's product. Instead of rejecting a change as soon as an in-house panel detects a significant difference, one can use a more lenient approach where the predicted consumer perception now plays a role in the final decision. This in turn should decrease the proportion of erroneous decisions, which can be categorized as type I errors, wrongfully concluding that consumers would be able to detect a sensory difference. A lower number of cost-saving opportunities would then be missed, which is particularly appealing in the ever more competitive food and personal care products industries.

As a last note, one needs to be aware that the findings presented in this paper are limited in scope as the numerical values cannot be readily generalized to other types of products, panels or testing conditions. Specific projects will need to be designed to generate those links for particular products and different testing conditions. However, the current research illustrates an approach which could have broad applicability to the solving of issues recurrently encountered by the sensory scientist. It should prove a valuable tool once the appropriate relationship has been determined.

References

Bende, M., & Nordin, S. (1997). Perceptual learning in olfaction: professional wine tasters versus controls. *Physiology and Behavior, 62*, 1065–1070.

Braun, V., Rogeaux, M., Schneid, N., O'Mahony, M., & Rousseau, B. (2004). Corroborating the 2-AFC and 2-AC Thurstonian models using both a model system and sparkling water. *Food Quality and Preference, 15*, 501–507.

Cardello, A. V., Muller, O., Kapsalis, J. G., Segars, R. A., Sawyer, F. M., Murphy, C., et al. (1982). Perception of texture by trained and consumer panelists. *Journal of Food Science, 47*, 1186–1197.

Chollet, S., & Valentin, D. (2001). Impact of training on beer flavor perception and description: are trained and untrained subjects really different? *Journal of Sensory Studies, 16*, 601–618.

Chollet, S., Valentin, D., & Abdi, H. (1997). Do trained assessors generalize their knowledge to new stimuli? *Food Quality and Preference, 16*, 13–23.

Dugle, J. (1997). Note on "experts versus consumers: a comparison". *Journal of Sensory Studies, 12*, 147–153.

Ennis, D. M. (1993). The power of sensory discrimination methods. *Journal of Sensory Studies, 8*, 353–370.

Ennis, D. M. (2001). Personal communication.

Ennis, D. M., Palen, J., & Mullen, K. (1988). A multidimensional stochastic theory of similarity. *Journal of Mathematical Psychology, 32*, 449–465.

Frijters, J. E. R. (1980). Three-stimulus procedures in olfactory psychophysics: an experimental comparison of Thurstone-Ura and three-alternative forced-choice models of signal detection theory. *Perception and Psychophysics, 28*, 390–397.

Gonzàlez, R., Benedito, J., Carcel, J. A., & Mulet, A. (2001). Cheese hardness assessment by experts and untrained judges. *Journal of Sensory Studies, 16*, 277–285.

Gonzàlez, R., Sifre, S., Benedito, J., & Noguès, V. (2002). Comparison of electromyographic pattern of sensory experts and untrained subjects during chewing of Mahon cheese. *Journal of Dairy Research, 69*, 151–161.

Hautus, M. J., & Irwin, R. J. (1995). Two models for estimating the discriminability of foods and beverages. *Journal of Sensory Studies, 10*, 203–215.

Hough, G. (1998). Experts versus consumers: a critique. *Journal of Sensory Studies, 13*, 285–289.

Juslin, P., & Olsson, H. (1997). Thurstonian and Brunswikian origins of uncertainty in judgment: a sampling model of confidence in sensory discrimination. *Psychological review, 104*, 344–366.

Lesschaeve, I., & Issanchou, S. (1995). Olfactory memory performance effects of sensory experience, familiarity and label use consistency. *Chemical Senses, 20*, 114–115.

Marchisano, C., Lim, J., Cho, H., Suh, D., Jeon, S., Kim, K., et al. (2003). Consumers report preferences when they should not: a cross-cultural study. *Journal of Sensory Studies, 18*, 465–486.

Mioche, L., & Martin, J. F. (1998). Training and sensory judgement effects on mastication as measured by EMG. *Journal of Food Science, 63*, 1–5.

Moskowitz, H. (1996). Experts versus consumers: a comparison. *Journal of Sensory Studies, 11*, 19–37.

Moskowitz, H. (1997). A reply: note on experts versus consumers: A comparison. *Journal of Sensory Studies, 12*, 155–168.

Moskowitz, H. (1998). Consumers versus experts in the light of psychophysics: a reply to Hough. *Journal of Sensory Studies*, 13.

Muñoz, A. M., & Chambers, E. IV, (1993). Relating sensory measurements to consumer acceptance of meat products. *Food Technology, 47*, 128–131, 134.

Muñoz, A. M., Chambers, E., IV, & Hummer, S. (1996). A multifaced category research study: how to understand a product category and its consumer responses. *Journal of Sensory Studies, 11*, 26–294.

O'Mahony, M. (1995). Sensory measurement in food science—fitting methods to goals. *Food Technology, 49*, 72–82.

O'Mahony, M., & Rousseau, B. (2003). Discrimination testing: a few ideas, old and new. *Food Quality and Preference, 14*, 157–164.

O'Mahony, M., Masuoka, S., & Ishii, R. (1994). A theoretical note on difference tests: models, paradoxes and cognitive strategies. *Journal of Sensory Studies, 9*, 247–272.

Parr, W. V., Heatherbell, D., & White, K. G. (2002). Demystifying wine expertise: olfactory threshold, perceptual skill and semantic memory in expert and novice wine judges. *Chemical Senses, 27*, 747–755.

Rousseau, B. (2001). The β-strategy: an alternative and powerful cognitive strategy when performing sensory discrimination tests. *Journal of Sensory Studies, 16*, 301–318.

Rousseau, B., & O'Mahony, M. (1997). Sensory difference tests: Thurstonian and SSA predictions for vanilla flavored yogurts. *Journal of Sensory Studies, 12*, 127–146.

Rousseau, B., & O'Mahony, M. (2000). Investigation of the effect of within-trial retasting and comparison of the dual-pair, same-different and triangle paradigms. *Food Quality and Preference, 11*, 457–464.

Solomon, G. (1990). Psychology of novice and expert wine talk. *American Journal of Psychology, 105*, 495–517.

Stillman, J. A. (1993). Response selection, sensitivity, and taste-test performance. *Perception and Psychophysics, 54*, 190–194.

Stone, H., & Sidel, J. L. (1993). *Sensory evaluation practices*. Orlando: Academic Press.

Thurstone, L. L. (1927). A law of comparative judgment. *Psychological Review, 34*, 273–286.