

STIMULATING CONSUMERS' AFFECTIVE MINDSET WHEN IMPLEMENTING HUMAN APPRAISAL SURVEYS

Julien DELARUE^a and Isabelle BOUTROLLE^b

^a *AgroParisTech, UMR 1145, France*

^b *Danone Research, France*

ABSTRACT

Hedonic tests classically implemented in human appraisal surveys are often criticized for the lack of commitment of the participants during the test and the artificiality of the hedonic response. The purpose of our work is to improve the way these tests can be implemented, especially when they are to be set up under controlled settings. In a first experiment achieved with 240 women, we tested whether an authenticity test could be used as an alternative to a traditional hedonic test when evaluating two brands of salted crackers. The principle is to induce a critical affectively negative attitude in the participants. This is achieved by telling them prior to the evaluation that the firm can raise its profit by making low-cost versions of their favourite product while selling it at the same price. The participants were then asked to say whether each presented sample was a genuine one or a copy. Results show that overall authenticity responses are closely linked to preferences but are more discriminant. Besides, the paired authenticity/preference data provide insightful information on the diversity in consumer perception. In a second experiment, we tested the possibility to induce contextualized responses by simulating context with short audio scenarios. In this study, we measured the appropriateness of a series of six imaginary eating situations for each tested product. Only one scenario allowed us to reveal a significant preference. This reveals that such a technique may be successfully implemented but needs careful preparation and prior understanding of consumers' eating habits.

Keywords: *hedonic testing, context, commitment, authenticity, scenarios*

Corresponding author: Laboratoire de Perception Sensorielle et Sensométrie, AgroParisTech, 1 avenue des olympiades, F-91300 Massy, France, julien.delarue@agroparistech.fr.

1. INTRODUCTION

Products developers in the food industry frequently check the performance of their products with the help of quantitative market research companies who usually propose the implementation of hedonic blind tests with large consumer panels. These tests are very frequently conducted in central facilities. Such Central Location Tests (CLT) consist of an evaluation of the products by the consumers after a brief exposure under standardized conditions of consumption. The main criticism raised against CLT is that the tasting conditions are very different from natural eating situations and do not allow the respondents to be emotionally involved like in actual food consumption, while making them often consider aspects they would never take into account when consuming normally [1, 2]. Hence, there is every chance that by inducing an analytical attitude in the subjects, these tasting conditions will bring about artificial responses. This effect is probably reinforced by the neutral, almost aseptic, context in which the tests are usually conducted. Various attempts to incorporate more realistic tasting conditions in CLT designs have succeeded in integrating specific contextual factors in consumer test designs. However, this reductive approach (singling out contextual factors) presupposes determination of the most natural consumption setting for the tested food product. However, food products are rarely consumed in only one kind of eating situation, which complicates the prediction of 'real life' hedonic responses based on hedonic data obtained from one-situation CLTs. Perceptual situations are indeed not exclusively defined by objective criteria; they are also defined by the subjects' conscious and subconscious intentions. Everyday life is a world of meanings rather than one of objective facts.

In a past experiment conducted with two brands of salted cheese crackers we observed no significant difference of liking using a CLT design, which was not the case in more realistic testing conditions when consumers consumed and evaluated the products at home [3]. At the time, we hypothesized that the non significant difference in liking revealed by the CLT was due to low involvement of the participants and to an analytical attitude during the hedonic test. Thus, we were interested in comparing again these two products with a more involving test protocol and in an attempt to trigger more affective responses. In a first experiment [4] we tested whether an authenticity test could be used as an alternative to a traditional hedonic test. The principle of this test is to induce a critical affectively negative attitude in the participants by telling them an upsetting story related to the products they are about to evaluate [5]. In a second experiment, we tested the possibility to induce contextualized responses with a situation-oriented approach by simulating context with short audio scenarios. In this study, we measured the appropriateness of a series of imaginary eating situations for each tested product.

2. EXPERIMENT 1: TEST OF AUTHENTICITY

2.1. Principle

The aim of this experiment was to test whether an 'authenticity test' could be used as an alternative to a traditional hedonic test. The principle of such a test, which was initially developed for difference testing by Mojet and Köster [5, 6], is to induce a critical affectively

negative attitude in the participants. This is achieved by telling them prior to the evaluation that the firm can raise its profit by making low-cost versions of their favorite product while selling it at the same price. Participants are then monadically presented with a series of samples consisting of multiple presentations of each of the versions to be compared and are asked to quickly taste them and to say for each of them whether they feel it is a genuine old one or a new copy. It is not necessary that the respondents have a reference for the products, but it is assumed that the one they like best is the one they pick as being original.

2.2. Materials and Methods

Two hundred and forty four women between 18 and 65 year old participated in the experiment (age 41.8 ± 12.8). They were recruited in three different French cities and were invited to a testing room to participate in a 15-min face-to-face interview. Two brands of salted cheese crackers, a national brand (indicated here as product X) and a private label product (indicated as product Y) were tasted and evaluated according to the authenticity test design described below.

In order to induce a negative attitude, the interviewer first told the subject the following upsetting story: *“We want to investigate the possibility of replacing the original version of our salted crackers brand by a copy manufactured with low-cost ingredients. This will only be possible if the consumers accept the low-cost product and if it tastes like the original version. Of course, the retail price will be the same, even though the ingredients are cheaper”*. The story was read aloud to the participants who also had a written version of the text in front of them during the task. After they had been told the story, the participants were asked to sequentially taste each item of a series of 13 crackers samples. They were told that the series would consist of ‘original’ samples and ‘copy’ samples and that several copy versions would be presented. For each tasted item, they were asked to indicate whether they felt it to be either an ‘original’ or a ‘copy’. Actually, apart from the first sample which was randomly chosen as X or Y, the next twelve samples consisted of an equal number of X and Y whose presentation sequence was balanced over the panel. Finally, at the end of the session, subjects were sequentially presented with a new pair of X and Y samples while specifying to them that one of these two samples was the original version and the other was a copy. Here, the subjects were asked to say which was their preferred sample.

2.3. Results

It seemed that the negative priming worked well for the majority of participants with some very negative reactions (one woman refused to participate in the study once she heard the upsetting story). However, participants were sometimes happy that, for once, we “told them the truth”. During the authenticity task, the repeated tasting was not really a problem, even if we observed of course some lassitude at the end of the sequence.

Overall data

One of the two products (product X) was significantly more often indicated as being ‘original’ than product Y. The test indeed invoked a higher percentage of ‘original’ responses to product X (42%) than to product Y (36%). The analysis of the overall data set according to the signal detection theory (SDT) shows that the mean of the absolute individual d' values

is higher than 0.5 (mean $|d'| = 0.94$) which indicates that the products are being discriminated. A logistic regression also shows that the product type, X or Y, brings a significant quantity of information ($p = 0.001$) to explain the variability of the response.

Sub-group analysis

A closer look at the individual d' shows an important disparity between individual values. This may explain why in spite of a significant overall discrimination, the proportions of 'original' responses for each product were very similar. Some participants used the product X as their 'original' reference whereas others rather referred to the product Y. A segmentation of the subjects on the basis of this distinction is thus discussed below.

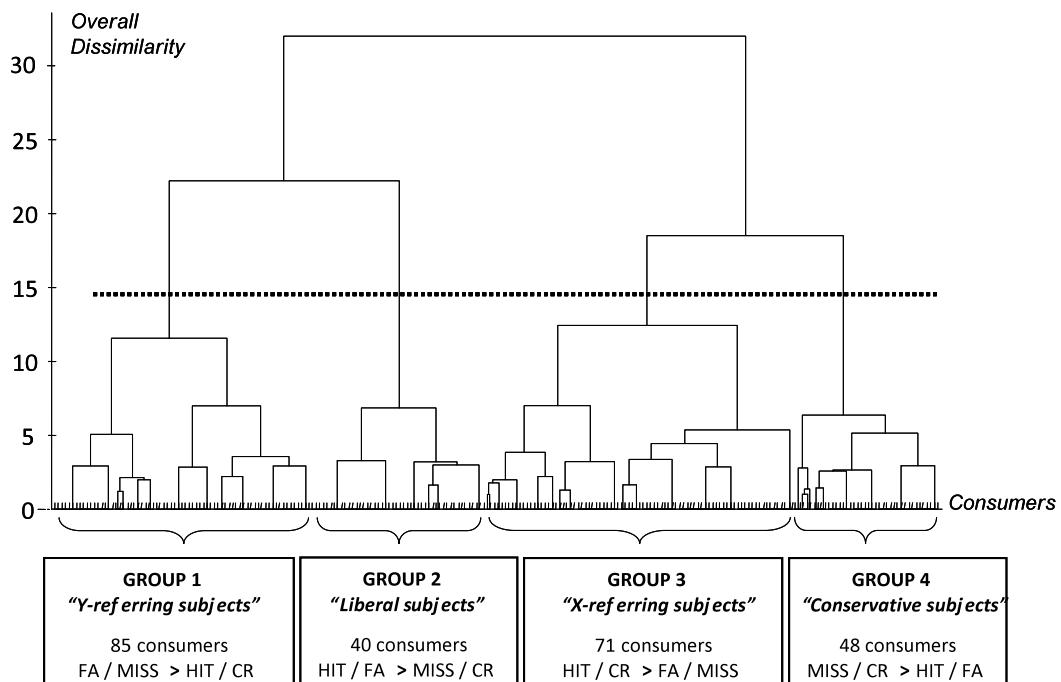


Figure 1: Cluster analysis of authenticity results.

Consumer groups and associated response patterns are summarized according to the SDT terminology (FA = False alarms, HIT = Hits, MISS = Miss, CR = Correct rejection).

Figure 1 shows the result of the cluster analysis applied to the SDT matrix of subjects' individual responses. This classification leads to four different groups of consumers. Group 1 (85 subjects) consists of respondents who tended to refer to the product Y as the guessed original product. In contrast, the response pattern of the consumers in group 3 (71 subjects) indicates that these subjects rather referred to the product X as being the original product. Two other groups differ in the decision criterion (τ) they used, group 4 (mean $\tau = 0.95$, SD = 0.19) using a more conservative (saying more often 'copy') one than group 2 (mean $\tau = -0.32$, SD = 0.28) which used a more liberal (saying more often 'original') criterion. Figure 2 presents the proportions of 'original' and 'copy' responses for each group according to the presented product. As can be seen, the subjects who referred to the product X as the original product are more discriminating than the subjects who referred to the product Y as the authentic one. Thus, the product X is more frequently perceived to be authentic for the "X-referring subjects" than is the product Y for the "Y-referring subjects".

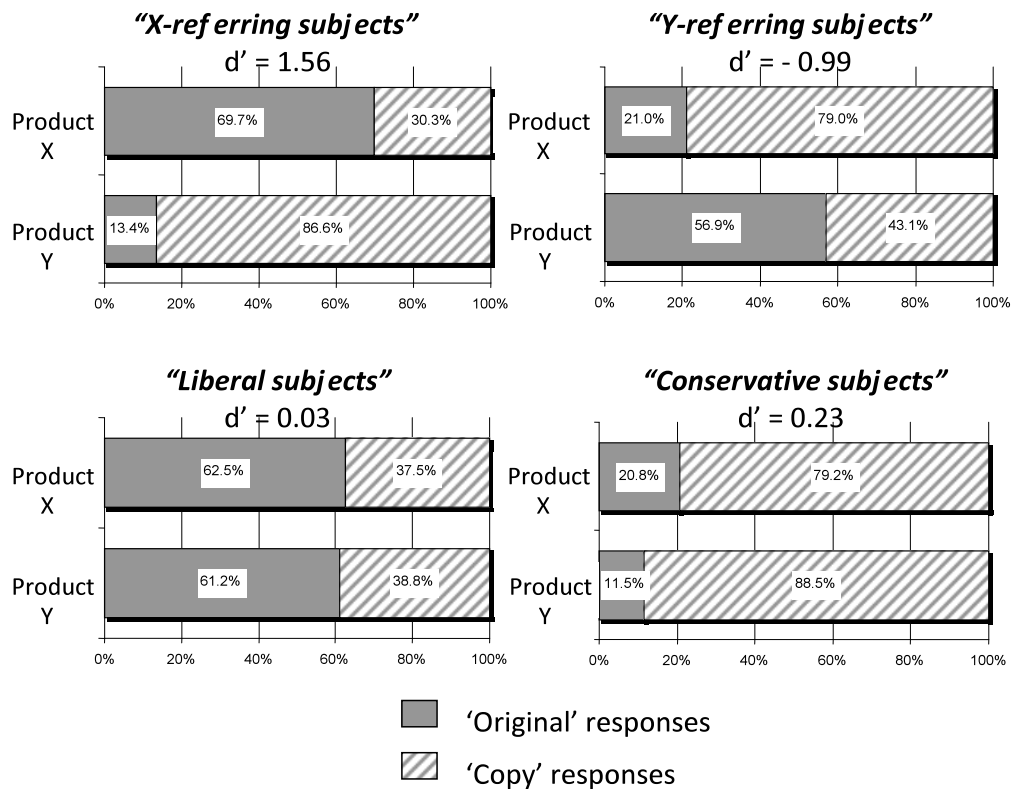


Figure 2: Percentages of 'original' and 'copy' responses according to the tasted sample for each consumer group.

It is worth noting that overall one third of the interviewed subjects responded without any authenticity tendency for one of the two products. However this does not necessarily mean that they do not perceive a difference between the products. In fact, it might also reflect a certain lack of confidence of the subjects while performing the task, resulting either in casual or in over-suspicious response behaviours.

Preference results

The preference question asked at the end of the session showed that, among the subjects who reported a preference toward one product, significantly more subjects ($p = 0.015$) stated a preference for the product X (57.6%) than for the product Y (42.4%). The preference question of the previously conducted traditional CLT (with the same products and a comparable consumer sample) did not reveal such a high preference for one product. This higher discrimination observed in the present experiment is promising, should one wish to reveal slight preferences.

Several explanations could account for the better performance of the combined authenticity-preference test at revealing significant preferences compared to that of the traditional CLT protocol. First, the specific repeated exposure design gives access to judgments (either authenticity or preference) that are based on repeated consumption of the tested products. On the contrary, during the traditional CLT, the subjects eat only a few crackers (and rarely more than 1 unit per product type) before stating their preference. Secondly, it is very likely that the design of the protocol affects the subjects' internal state

and thus the form of their involvement (intuitive rather than analytical) in the evaluation task. Eventually, we observed that the subjects were very concentrated on the task during the authenticity-preference test, which was less apparent during the standard CLT previously implemented. We may thus hypothesize that the authenticity task increased the subjects' involvement during the whole task until the end of the test including the preference question.

Authenticity and preference

The general response pattern observed for the authenticity task and for the preference question suggests that the authenticity is indeed a good "hidden" liking question. At the group level, a convergence between the preferences and the perceived genuineness of the products was found since the product X is the preferred product and is also considered to be the most authentic product. However, on an individual basis, it seems that for some subjects, the sense of authenticity inferred from the sensory perception of the products can be different from liking. Hence, we believe that this kind of paired authenticity/preference data brings some insightful information, notably providing insights in the diversity of consumer perception. Actually, food developers might be interested in both the liking and the perceived genuineness of their recipes, especially when the results are not correlated. For example this approach could be well appropriated when investigating the links between consumer judgment and choice behavior with variables such as brand or price. Implementation of this authenticity task to test innovative products could also be very informative, by measuring sensory expectations conveyed by the innovative concept. In such a task, consumers would have to guess the authenticity of formulated recipes that they might never have experienced before.

3. EXPERIMENT 2: CONTEXT SIMULATION WITH AUDIO SCENARIOS

3.1. Principle

The aim of this experiment was to measure the fittingness of food products to situations, as had already been proposed by Schutz et al. (1977) in the 'item by use appropriateness' method [7]. That method consisted in presenting the subjects with a list of foods and list of possible uses (time of day, site, occasion, physiological state, person, etc.) and having the subjects score the appropriateness of each food product for the set of uses proposed. However, the few studies that address the appropriateness of a product for a situation, described the corresponding situations only in a few words (e.g.: 'when I'm eating in front of the television'). That approach to the question is open to criticism since it does not necessarily enable the subject to become involved and really 'project' him/herself into the situation in question. One way of improving situation appropriation by the subject is to make use of his/her autobiographic memory [8]. The subject is thus considered a historical being endowed with an autobiographic memory composed of events that have been personally experienced, located and dated, and are specific to each individual [9, 10]. Here we attempted to evoke and prime imaginary situations with the help of auditory scenarios, as suggested by Köster [11].

3.2. Materials and Methods

The efficacy of the approach was tested on 240 women by collecting consumption intentions in various situations evoked through audio scenarios with a view to comparing the sensory performances of two types of salty cheese-flavored crackers. While very similar, the two products nonetheless differed in terms of appearance, taste and texture. Our aim was to verify whether the sensory differences conditioned a different consumption intent in 6 consumption situations: week-day drinks before the meal, weekend drinks before the meal, picnic, snacking while relaxing, snacking while working, and snacking during transport. We decided to focus on these 6 situations after having conducted a series of 20 individual interviews with consumers.

The various scenarios were compiled in the first person singular in a relatively simple style and gave information on a certain number of contextual variables such as the physical environment, but also on the subject's interior condition [12]. For example, the scenario used to illustrate the weekend drinks before the meal situation was as follows: *'It's the long awaited time for drinks before dinner at the weekend. I can at last spend some time with my loved ones. As usual, the conversation is lively and drinks last longer than expected. After a few drinks, it's probably time to have something to eat but nobody seems to want to sit down to the meal. Despite the big meal that's on the way, I take a dish of crackers and offer them around without forgetting to help myself'*. Each subject thus listened to six scenarios one after the other and formulated his/her intent to eat the tested cracker in the situation (from 1 'no, certainly not' to 10 'yes, absolutely') together with the frequency of that situation in everyday life (from 1 'never' to 10 'very regularly').

3.3. Results

As observed in Figure 3, the intent to consume recipe Y was higher than that for recipe X in almost all situations, but the difference was only significant for the weekend drinks before the meal scenario. Moreover, the intent-to-consume scores for the two products were significantly higher for the weekend drinks than for the other situations (t-test: $t=2.04$, $p < 0.05$). Among the remaining five situations, the products obtained significantly higher scores in the week-day drinks before the meal scenario and in the picnic scenario than in the three snacking scenarios.

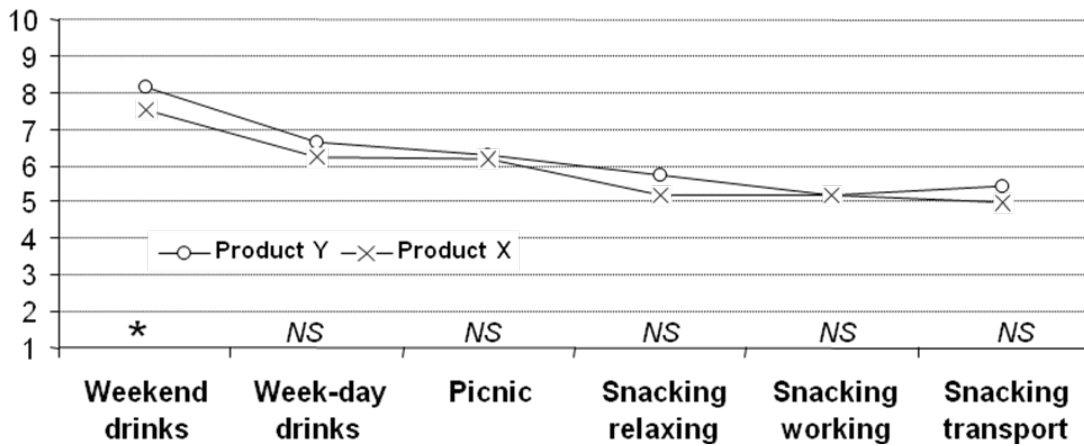


Figure 3: Mean intent to eat scores for the two cracker products by situation.
t-test results: NS: not significant; *: p-value < 5%.

From those results, an overall intent to consume index (individual mean of the six intent to consume scores weighted by the frequency of those situations in the subject's lifestyle) was generated. Rather than a simple mean intent to consume, the index has the advantage of giving, for each subject, more weight to the intents to consume in very frequent situations than the intents in infrequent situations. For the panel as a whole, an overall mean intent to consume score was obtained for recipe X (mean weighted intent score = 6.66) and recipe Y (mean weighted intent score = 7.16). The difference between the two scores was significant (non-paired t-test: $t=2.10$, $p = 0.037$). This strengthens the superiority of the overall intent to consume for recipe Y vs. recipe X. The intent-to-consume by situation data show very high standard deviations reflecting the diversity of the participants with regard to the intent to eat crackers in the various situations. This affords the possibility of investigating the data by subject group on the basis of the responses formulated with the situation occurrence frequencies in the participants' everyday lives as indicated by the participants.

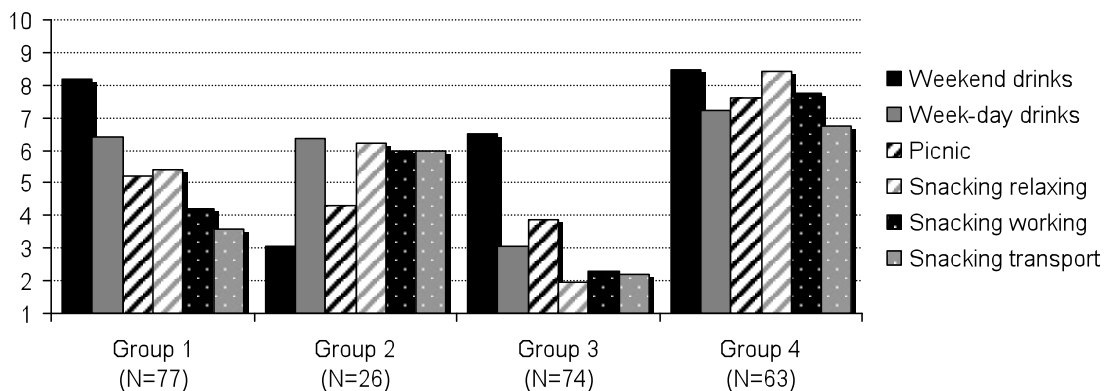


Figure 4: Mean situation frequency profiles for the four groups of subjects identified by k-means classification.

Four different cracker consumption profiles were thus determined by the use of k-means classification and illustrated on Figure 4. Group 1 consists of subjects who prefer, for the

food type in question, consumption situations associated with a social event (drinks before the meal on weekends and picnics) or a solitary but relaxing event (drink before the meal during the week or relaxed snacking). The situations in which consumption is more utilitarian or functional (work, transport) are less associated with the food category. Group 2, which was smaller, consisted in the consumers who do not eat crackers in an environment with marked social interaction (weekend drinks before the meal and picnics) but only when alone (drinks before the meal during the week and snacking). Group 3 consists of the subjects who mainly eat the food type during drinks before the meal at the weekend. In contrast, group 4 consists of participants who eat crackers regularly in all of the proposed situations. These results show that product use is to be taken into account not only in study design but also in the interpretation of the data generated. A real consumer typology based on food eating habits is observed. An understanding of the typology thus seems essential in order to elucidate consumer preferences.

4. CONCLUSION

The first experiment reveals that the implementation of an authenticity test design is feasible and has many advantages. First, the repeated exposure design provides judgments that are based on several consumptions of the tested products. To this end, this was a very satisfactory procedure, since it is usually difficult to drive participants into eating several samples repeatedly without explicitly forcing them. This may lead to a better external validity than that of hedonic data collected from first impression only. This can be related to Berlyne's categorization of hedonic task response behaviours into either specific or diversive exploration [13]. Specific exploration is not concerned with pleasure but with resolving puzzling stimulation and reducing uncertainty. Thus, instead of real preference or liking, when they test first impression, market researchers obtain responses that are mostly based on curiosity and on the desire to learn more about the products. The diversive exploration focuses on the pleasurable aspect of the product and thus for the more durable appreciation of it. This is only possible after specific exploration is completed and the uncertainties are resolved. Besides, the variation in tested product consumption intents, depending on the situation, observed in the second study points to the value of taking into account food product use in consumer tests. This observation could be made even though the autobiographical priming only had a minor effect on the hedonic response.

REFERENCES

1. Köster E.P., Diversity in the determinants of food choice: A psychological perspective, *Food Quality and Preference*, Vol. 20, pp. 70-82, 2009.
2. King S.C., Weber A.J., Meiselman H.L., and Lv N., The effect of meal situation, social interaction, physical environment and choice on food acceptability, *Food Quality and Preference*, Vol. 15, No 7-8, pp. 645-653, 2004.
3. Boutrolle I., Delarue J., Arranz D., Rogeaux M., and Koster E.P., Central location test vs. home use test: Contrasting results depending on product type, *Food Quality and Preference*, Vol. 18, No 3, pp. 490-499, 2007.

4. Boutrolle I., Delarue J., Köster E.P., Arranz D., and Danzart M., Use of a test of perceived authenticity to trigger affective responses when testing food, *Food Quality and Preference*, Vol. 20, No 6, pp. 418-426, 2009.
5. Mojet J. and Köster E.P., Onderzoek naar de waardering van drie laag alcoholische bieren (Research on the appreciation of three low-alcohol beers). *Report Psychological Laboratory, Utrecht University*, 1986.
6. Köster E.P., Les épreuves hédoniques, In *Evaluation sensorielle, manuel méthodologique (2nd ed.)* (Eds F. Depledé and R.K. Striegler), Lavoisier: Paris, pp. 182-206, 1998.
7. Schutz H.G., Moore S.M., and Rucker M.H., Predicting food purchase and use by multivariate attitudinal analysis, *Food Technology*, Vol. 29, pp. 50-64, 1977.
8. West P., *Reproducing naturally occurring stories: vignettes in survey research*. 1982, MRC Medical Sociology Unit: Aberdeen.
9. Drifford V., Lemmet J., Bernard P., Laporte V., and Chambres P., Acceptance tests and memory of the context of consumption: analysis and experimentation, *Viandes et Produits Carnés*, Vol. 16, pp. 107-111, 1995.
10. Jaeger S.R. and Meiselman H.L., Perceptions of meal convenience: the case of at home evening meals, *Appetite*, Vol. 42, pp. 317-325, 2004.
11. Köster E.P., The psychology of food choice: some often encountered fallacies, *Food Quality and Preference*, Vol. 14, pp. 359-373, 2003.
12. Krosnick J., Response strategies for coping with the cognitive demands of attitude measures in surveys, *Applied Cognitive Psychology*, Vol. 5, pp. 213-236, 1991.
13. Berlyne D.E., *Conflict, arousal and curiosity*. 1960, New York: McGraw-Hill.