"Who will buy it?": Analysis of package design of organic soy sauces in view of users' lifestyle

Using data mining techniques

Yoshie Kiritani*, Hisako Harada**, Kenji Ota***

* Chiba University, kiritani@faculty.chiba-u.jp ** Chiba Universit, tttttahiti@gmail.com

Abstract: The purpose of the present study was to reveal the relation between users' properties like cooking and buying behavior and their preference for the packages, using data mining techniques. A Hayashi's quantification theory III was done for the 44 samples of organic soy sauce in the current market, to make clear the features of the current products. A cluster analysis with Ward method revealed 8 clusters: 'literal expression', 'decoration', 'illustration', 'green', 'digital processing', 'warm color and simple', 'decorative frame', and 'much information'. Fifty-six adults in their 20s to 60s answered about the life style, looked at the real bottles of 8 samples representing each cluster and judged whether they wanted to buy them. As a result, the purchase frequency of organic foods did not correlate with the willingness to buy. A decision tree analysis and a market basket analysis revealed rules to show who wanted to buy the samples. Frequently cooker for themselves or for cohabiter(s) wanted to buy the moderately popular samples. It was the package using green color as a code color of organic foods that was favored by people who pay attention to package design. The present study revealed 6 suggestions for new design of organic soy sauces.

Key words: user behavior, psychology of design, graphic design, communication, data mining techniques

1. Introduction

The present study is about the package of an organic food, because it would play an important role according to the following story.

Kubota (2003) [1] insisted that, in Japan, organic agricultural movement came from reflections and criticism against the modern agriculture. Before 1960's, Japanese farm families had been self-contained. Their life had been sustainable and stable, having a small number of farm animals and utilizing mountain forests. The Agricultural Basic Act (Act No. 127 of 1961) [2] encouraged the use of a large amount of chemical fertilizer and agricultural chemicals. Basic policies were expansion, mechanization, and monocropping, and the production efficiency was emphasized at that time. This institutionalized agriculture has lead better productivity and, at the same time, caused pollution and environmental issues. The organic agriculture was a citizen action against the institutionalized agriculture.

In the mid-1980, channels of distribution became diverse. The more the number of retailing stores dealing with the organic foods increased, the more producers grew apart from consumers. They relied on the indication of 'organic', but irresponsible marks flooded [3]. The Act for Standardization and Proper Labeling of Agricultural

^{***} Chiba Universit, ota.kenji.21@gmail.com

and Forestry Products established in 2000 stipulates third-party certificate for the producers to display 'organic'. The certificated producers can put the organic Japanese Agricultural Standards mark on merchandize. The law forbids use of not only the mark but also the word of 'organic' or confusable words without the certification (Ministry of Agriculture, Forestry and Fisheries website). However, in fact, over half of respondents did not know these rules of the organic JAS mark [4].

A multiple channel of distribution of the organic foods provides us many opportunities to purchase the organic foods. But 78.8 % of respondents answered that they bought the organic foods in supermarket [4]. This tendency can count against the organic foods from the point of market principles, because these foods are usually more expensive than the ordinary. To compete in the market, the organic foods should use the design tactics. Now, the package becomes important.

For periodic purchasers, the package of organic foods may not be so important, because these people are compelled to buy these foods for their or family's health or really understand the value of organic foods. However, in supermarket, there are the persons who are on the edge discriminating the consumers of the organic foods from others who are not interested in them. If they find some values in the organic foods, they will buy them. The package will be an effective factor in the field of purchase. What kind of people focus on the package design? What kind of design will the people prefer rather than others?

We selected soy sauce as an organic food to research this issue, because a preliminary survey revealed that seasonings have mass appeal to beginners. Soy sauce has board utility in cooking of Japanese meal, comparing to miso (bean paste).

The purpose of the present study was to reveal the relation between users' properties like cooking and buying behavior and their preference for the packages. Data mining techniques would reveal the relation between types of persons and preferred packages; we could know which kinds of people would buy a certain package. If we found tendencies of packages to attract attention, we would change the design efficiently.

2. Features of market products

2.1 Grouping of packages of soy sauce

2.1.1 Samples and categories

Forty-four packages of organic soy sauce and 45 packages of regular soy sauce were collected. These products were relatively easily-available around Tokyo. The number of regular soy sauce was matched to that of the organic soy sauce, because countless products were sold in Japan.

The samples of organic soy sauce were the real things, so that their both sides of package were checked. Most of samples of regular soy sauce were pictures downloaded from the site. Some of them were the real things, not just pictures, if there was information about producer or production on the reverse side of package.

Items to describe the package were determined by reference to Ito, Kiritani, Ohara, Tamagaki and Miyazaki (2007) [5], which were about graphic, typography, color, layout and bottle. Seventy-two categories, in total, were set under these items.

2.1.2 Results

A Hayashi's quantification theory III analyzed all samples and derived two axes. The first axis has 0.44 of eigenvalue and 0.66 of coefficient correlation value; these values for the second axis were 0.31 and 0.55, respectively. Categories in the positive part of the first axis were 'Gold or silver' of the code color, 'partial decoration' of package, 'under 2 colors' of the color, 'same' of the alignment of product name and graphics, and so on. These categories suggested high-class atmosphere. On the other hand, representative categories appeared in the negative part of the first axis were 'decorative frame' of graphics and 'Mincho font' that were usually and traditionally used in the popular soy sauces. Thus, the first axis was considered as 'special versus daily-use'. Categories in the positive part of the second axis were 'simple' of the layout, 'bland-name logos' of the graphic and so on, which suggested simple configuration using symbolic figures. Categories in the negative part of the second axis were 'complex' of the coloration, 'decorative frame' of the graphics, 'oblique or curve' of the typography orientation. Thus, the second axis was regarded as 'simple versus complex'.

Most of samples of organic soy sauce are encircled, which locate in the positive area of the first axis. Thus, the packages of organic soy sauce were regarded as special. From the 'simple-complex' viewpoint, the sample of organic soy sauce had no features, because they were near the center of the second axis.

Differences of packaging between the organic soy sauce and the regular soy sauce are listed in Table 1. It represents the features of majorities of both soy sauces.

Table 1. Difference in packages between organic soy sauce and regular soy sauce

Visual elements	Organic soy sauces	Regular soy sauces	
Graphics	No graphics (if using, geometrical	Decorative frames, illustration of	
	configurations, illustration of raw	raw materials, and bland name	
	materials or production method)	logos	
Letters used in product name	Hairy brushes	Digital fonts	
Orientation of typography	Vertical	Horizontal	
Color	Under 3 colors/ lower chroma with	Over 4 colors	
	other few colors (if higher, without		
	other colors)		
Material of package	Glass bottle	Pet bottle	

2.2 Grouping of packages of organic soy sauce

2.2.1 Samples and categories

Another analysis of Hayashi's quantification theory III was done for the same 44 samples of organic soy sauce, to make more clear the features of the current products. Some new categories, for example about size of organic label, were added for the analysis.

2.2.2 Results

Two axes were adopted. The eigenvalue of the first axis was 0.33 and its coefficient correlation value was 0.58. These values for the second axis were 0.29 and 0.54, respectively. The positive part of the first axis had categories like 'graphic' of the production method, 'graphic' of the producers. Its negative part had categories like 'no information' of the production method and 'no information' of the producers. Thus, the first axis was regarded as 'informative versus less informative'. In the positive part of the second axis, there were categories like 'dynamic' and 'mixture' of the rhythm of product names. On the other hand, in the negative part of the second axis, 'stylized letters' of the font of product names and 'horizontal' of the typography orientation. Thus, the second axis was called as 'hand-made versus digital'.

A cluster analysis with Ward method revealed eight clusters of 44 samples (see appendix). These clusters were named as 'literal expression' (A), 'decoration' (B), 'illustration' (C), 'green' (D), 'digital processing' (E), 'warm color and simple' (F), 'decorative frame' (G), and 'much information' (H). The features could be described as follows and distinctive features of these groups that differentiate from each other were listed in Table 2.

Table 2. Differential features of 8 groups

	Name	Features	Rank	Support rate (%)	Value of Chi-square (Freedom =3)
A	Literal expression	Explanation put also in the front label in some cases	4	58.9	5.933
В	Decoration	Gold or silver as color Bagged bottle or with decorations	1	96.4	1.372
С	Illustration	Illustration of the production method or raw materials Geometric figures drawn by hairy brush	3	69.6	2.817
D	Green	Green as color	5	46.4	6.583
Е	Digital processing	Brand logo Horizontal composition of letters Color gradations in the background	6	37.5	0.121
F	Warm color and simple	Pet bottles	8	14.3	4.868
G	Decorative frame	Decorative frame common to the usual soy sauce Curve composition of letters	7	17.9	0.630
Н	Much information	Much information in small bottle	2	75.0	1.503

The samples in the 'literal expression' group were, literally, consisted of letters without graphical components. These letters were usually hairy brushes that were visually attractive and elegant. No graphics made enough space to put information of the product. The package of all samples was made of glass. There were 7 samples in this group.

The samples in the 'decoration' group were classified into two sub-groups. One was consisted of letter-oriented like the group of 'literal expression'. Another contained simple geometrical figures besides letters. For both of them, the number of design elements was small, so that the impression of package was sophisticated. Moreover, some of them were, partially or wholly, covered, or gold-silver stamping. They seemed luxury. The package of all samples was made of glass, like the group of 'literal expression'. There were 8 samples in this group.

The main feature of the 'illustration' group was illustration overlapping the product name. The names were hairy-brush. The most popular illustration was about the tools, actions, or raw materials for the production. The next was geometrical figures by hairy brushes. Warm colors, like red, yellow and beige, were usually used. Almost all of samples were made of glass. There were 11 samples in this group. It was the biggest of 8 groups of organic soy sauce.

The color green was used in the group of 'green' as a code color. Green was used in the package of soy sauce to represent healthy aspect like organic or law salt, especially in product lines of the major manufactures. This group contained pet bottles and carton as package. There were 6 samples in this group.

The group of 'digital processing' contained digital fonts or a gradational background like a computer graphics. Some illustrations were also put but were brand logos or motives of the brand, which were not directly related to the materials used in the soy sauce. There were 5 samples in this group.

Package of the group of 'warm color and simple' was pet bottle. They were plain and modest design. Yellow or orange were used. There were 4 samples in this group.

The group of 'decorative frame' adopted the decorative frame usually used in the package of soy sauce of the major manufactures. Many colors were used and coloration was complex. The orientation of typography was also multiple. These packages gave busy and strong impression. There were 2 samples in this group.

Just 1 sample constituted the 'much information' group. The information of production was expressed not only by letters but also by graphics.

3. Users' life style and their taste in packages

3.1 Method

3.1.1 Participants and samples

Fifty-six adults (18 males and 38 females) in 20s to 60s (31 of 20s, 8 of 30s, 14 of 40's, 1 of 50s, and 2 of 60s) participated in the evaluation of the organic soy sauce packages. These participants should be representative of the consumers in the real field. Then, we asked and selected the participants to get a diversity of interests in the organic foods. They were students of cooking classes of the organic foods and university students.

The samples to be evaluated were the above-mentioned representatives of 8 groups of the organic soy sauce. There were 8 stimuli: 'literal expression', 'decoration', 'illustration', 'green', 'digital processing', 'warm color and simple', 'decorative frame', and 'much information'. All of them were the real bottles of soy sauce, not just photos.

3.1.2 Procedure of evaluation

The participants picked up each stimulus and judged whether they wanted to buy it. The order of stimuli for evaluation was different from individual to individual. Before the evaluation, each participant answered about the life style, including demographic features. All of questions to measure the life style are presented in Table 3. Question 5, 6 and 9 were of multiple-choice. There was no time restriction for the evaluation and the answer.

Table 3. Questions used to determine the life style

	Questions	Choices		
1	Gender	Male/ Female		
2	Age	20s/ 30s/ 40s/ 50s /60s		
3	Frequency of cooking	Daily/ Sometimes/ Few/ None		
4	Frequency of cooking for cohabiter(s)	Daily/ Sometimes/ Few/ None or Single life		
5	Places to buy foods	Supermarket/ Department store/ Specialized store/ Home- delivery service/ Local retail store/ Convenience store		
6	Points of focus in food shopping (except for price and taste)	Package/ Information of raw materials/ Freshness or freshness date/ Brand, maker, or producer/ Nutrition labeling/ Sales copy		
7	Knowledge of recognition system of organic foods	Almost all of standards, mark and display rule/ Mark and display rule/ Only mark/ None		
8	Frequency of buying of organic foods	Periodical/ Sometimes/ Once or twice/ None or no memory		
9	Images of organic foods	Safety of foods/ Healthy/ Good taste/ Sustainable farming method/ Activation of farming community/ Nutritious/ Luxury item/ Difficult to get/ Dirty-looking/ Bad taste/ Fanatic/ Hypocritical		

3.2 Results of simple tally of the users' life style

About the frequency of cooking, 'Daily' was 53.6 %, 'Sometimes' 28.6 %, 'Few' 12.5 % and 'None' 5.4 %. This data means that for most of participants the cooking was a daily-life activity. On the other hand, as for the frequency of cooking for cohabiter(s), 'Daily' was 35.7 %, 'Sometimes' 14.3 %, 'Few' 7.1 % and 'None' 42.9 %. The participants could be divided into 2 groups: one of people who had family or roommates and another of people who lived alone. The former was of students of cooking classes and the latter was of university students.

The degree of knowledge of recognition system of organic foods was dispersed: 'Almost all of standards, mark and display rule' 17.9 %, 'Mark and display rule' 23.2 %, 'Only mark' 26.8 %, and 'None' 32.1 %. Over half of participants knew at least the mark of organic foods.

The frequencies of buying of organic foods were also dispersed: 'Periodical' 26.8 %, 'Sometimes' 17.9 %, 'Once or twice' 26.8 %, and 'None or no memory' 32.1 %. The researchers had intended to get a diversity of participants. The data successfully back up the intention.

As the images of organic foods, main images were 'Safety of foods' (94.6 %) and 'Healthy' (85.7 %). 'Good taste' (51.8 %), 'Nutritious' (51.8 %), and 'Luxury item' (53.6 %) were imaged by about the half of participants. 'Difficult to get' (23.2 %), 'Sustainable farming method' (12.5 %) and 'Activation of farming community' (12.5 %) were less common. Images of 'Dirty-looking' (8.9 %), 'Bad taste' (0.0 %), 'Fanatic' (1.8 %) and 'Hypocritical' (3.6 %) were very few. The general image of organic foods was good, although it was not translated into the understanding of agricultural issues.

As for the place to buy foods, over 80 % of participants used 'Supermarket'. 'Convenience stores', 'Homedelivery service', and 'Specialized stores' were used by about 30 % of participants. 'Department stores' and 'Markets' were usually not used.

A for the points of focus in food shopping except for price and taste, there was no choice getting overwhelming popularity. It was half of the participants (53.6 %) who pay attention to the package.

3.3 Results of simple tally of the package evaluation

The ranking of preference was presented in the righter part of Table 2. The most favorite package was 'decoration', 96.4 % of participants wanted to buy. On the other hand, 'warm color and simple' sample was the worst, which got favor by only 14.3 % of participants. The samples of 'literal expression', 'green', and 'digital processing' were moderate.

Cross tables revealed that there was no relation between users' purchase frequency of organic foods and the willingness to buy the samples. In the Table 4 there are also the values of Chi-square for tests of the independence. The significant level 0.05 for the 3-degree-of-freedom is 7.815, so that all of values in the Table 4 were not significant. Thus, regardless of the purchase frequency, the tendency of preference of package was constant; 'decoration' was the most favorite sample and 'warm color and simple' and 'decorative frame' were not popular in all levels of the purchase frequency.

3.4 Relation between users' properties and the willingness to buy the samples

The popular samples like 'decoration', 'much information' and 'illustration' got support over about 70 % of the participants. Thus, the present research did not analyze them any more. Since the majority of people would buy them, there is no point in finding the special feature of buyer. And there will be no problem, because everybody will willingly buy the merchandises with these packages. Thus, the present research examined which type of

people wanted to buy 'literal expression' sample, 'green' one, 'digital processing' one, 'decorative frame' one, and 'warm color and simple' one, using a decision tree analysis and a market basket analysis. Table 4 shows rules of purchase generated by decision tree analysis, and Table 5 is about those of non-purchase. Table 6 shows features of purchaser who were rare but surly buy certain samples from basket market analysis.

Table 4. Rules of purchase generated by a decision tree analysis

Sample	Rule (Properties of expected buyer)	Satisfaction
Literal expression	Cooking daily or sometimes for cohabiter(s) Not-cooking or cooking rarely for cohabiter(s) and attention to freshness or its date	82.1 % of 28 persons 56.3 % of 16 persons
Green	Cooking daily Cooking daily and Age of 20s, 30s or 40s	63.3 % of 30 persons 67.9 % of 28 persons
Digital processing	Having an image of difficulty to get Having an image of difficulty to get and of luxury item	69.0 % of 13 persons 89.0 % of 9 persons

Table 5. Rules of non-purchase generated by a decision tree analysis

Sample	Rule (Properties of expected buyer)	Satisfaction
Literal expression	Not-cooking or cooking rarely for cohabiter(s)	64.3 % of 28 persons
Green	Not-cooking, cooking rarely or sometimes	73.1 % of 26 persons
Digital processing	Not-having an image of difficulty to get	72.0 % of 43 persons

Table 6. Lower-frequent but sure properties of purchase revealed by a basket market analysis

Sample	Properties of expected buyer)	Support	Confidence	Lift
	Cooking daily and having an image of luxury item	0.161	1.0	1.70
Literal expression	Attention to information of raw materials and to nutrition labeling	0.161	1.0	1.70
	Attention to nutrition labeling and having an	0.161	1.0	1.70
	image of good taste Attention to freshness and its date and to nutrition labeling	0.161	1.0	1.70
	Cooking daily for cohabiter(s) and attention to	0.107	1.0	2.15
	nutrition labeling			
	Age of 40s and attention to package	0.107	1.0	2.15
Green	Age of 40s and attention to nutrition labeling	0.107	1.0	2.15
Green	Use of specialized store and attention to	0.107	1.0	2.15
	package			
	Use of local retail store and not-use of	0.107	1.0	2.15
	convenience store			
	Attention to sales copy and having an image of	0.107	1.0	2.67
Disital massassins	difficulty to get			
Digital processing	Cooking sometimes and having an image of nutriment	0.107	1.0	2.67
Decorative frame	Use of home-delivery service and having an	0.053	1.0	5.60
	image of activation of farming community			
	Cooking rarely and not-having an image of	0.053	0.6	4.20
	luxury item			
Warm color and	Cooking rarely and having an image of health	0.053	0.6	4.20
simple	Cooking rarely and not-having an image of	0.053	0.6	4.20
	activation of farming community			
	Cooking rarely and having an image of safety of foods	0.053	0.6	4.20

3.4.1 About 'literal expression' sample

Decision tree analysis got 87.5 % of discrimination rate. The first generated rule in Table 4 tells us that the person who cooks for her/himself or for cohabiter(s) will buy this sample. On the other hand, the person who usually does not cook for others will not but it (Table 5). However, even this kind of people will buy it, if they are sensitive to the freshness of food (the second rule in Table 4).

More concrete features of purchasers will be imaged by the result of a market basket analysis (Table 6). They would be the type of homemaker.

3.4.2 About 'green' sample

Decision tree analysis got 82.14 % of discrimination rate. The generated rules in Table 4 shows that cooker will buy it, as in the literal expression sample. Table 5 revealed that the person who has lower frequency of cooking would not but the sample. Market basket analysis revealed some concrete characters of the purchaser (Table 6), which were consistent with the rules by decision tree analysis. Age of 40s is also a key. The importance of the package design is also suggested for this sample.

3.4.3 About 'digital expression'

Decision tree analysis got 92.86 % of discrimination rate. The generated rules in Table 4 tell us that the person who regarded the organic foods as difficult to get tended to buy this sample. If this type of person imaged the organic foods as luxury items, the possibility of purchase increased. Table 5 tells us that the person who did not have an image of difficulty to get would not buy this sample.

Market basket analysis (Table 6) revealed 2 relatively rare but certain rules of purchase; one is 'attention to sales copy and image of difficulty to get', the other is 'cooking sometimes and image of nutriment'.

Thus, the key point of purchase of 'digital expression' sample was an image of organic foods as difficulty to get.

3.4.4 About 'decorative frame'

This sample finished second to last, as shown in Table 2. Although most people did not want to buy it, market basket analysis revealed a rare but certain rule of purchase (Table 6). If the person used home delivery services and considered the organic foods as contribution to activate farming communities, they would buy this sample. Thus, this sample is not liked by everybody but wins support from the persons with much knowledge about the organic foods.

3.4.5 About 'warm color and simple'

Decision tree analysis was not suitable, but Market basket analysis revealed who were the purchasers. The common feature of them was 'not-cooking'. Thus, this sample was the worst popular one and whose purchasers usually did not cook.

3.5 Discussion

The previous section revealed the lifestyles of supporters of each moderately popular sample of organic soy sauce. From these properties, we can concretely image some purchasers.

As the case of 'literal expression', the higher frequency of cooking was a point. The cooking can be not only for oneself but also for cohabiter(s). Attention to information about foods was also related. Thus, we can image

the person of dietary manager in the house like homemaker as a type of purchasers of this package. They can be of firm character, because they seem to pay attention to health of others. According to the features in the current packages in Table 2, in some cases, this type of package contains the literal information not only in the back face but also in the front. To get more attention of targets, these explanations should be put more.

Also for the case of 'green', the higher frequency of cooking was a point. Although, just like the case of 'literal expression', the homemaker of firm character was imaged, live-alones were also possible because of the results of the decision tree analysis. Usage of stores suggested their peculiar taste. The most important matter was that they pay attention to the food package at the time of purchase. The purchasers would be with peculiar taste and look at the package, so that the package of this green type should be elaborated. It will be worthy of designing. But, it should not change the color code, green, which is the visual prompt for the purchasers.

For the case of 'digital expression', appeal of difficulty of getting will be effective. Sales copy is also watched (Table 6). According to the features in the current packages in Table 2, the package does not originally look much like a soy sauce. Thus, it will be interesting for this group to dare a design that does not look like soy sauces.

On the other hand, it will be possible for the group of 'decorative frame' not to change any design in the package. There were certain purchasers, although they were a only few people. They must look at the decorative frame in the package design.

Finally, the results suggest the stop of design of 'warm color and simple', because it was not popular and whose purchasers seem to be less interested in cooking.

5. Conclusion

The present study revealed the popularity of the package of organic soy sauces and which types of people tended to buy the samples. The buying frequency of organic foods did not predict the willingness to buy. On the other hand, the frequency of cooking was an effective factor for the modest popular package.

'Decoration' type sample was the most popular one. This seems sophisticated or luxury package, because the number of design elements was small and some of them were, partially or wholly, covered, or gold-silver stamping. Although 'literal expression' sample, 'green' one, and 'digital processing' were modest about the popularity, they were surly supported by house keepers, by people with peculiar taste, and by people with interest in rarity, respectively. Images of the purchasers will lead to new designs. The sample of 'decorative frame' was not popular package but supported by people with much knowledge about the organic foods. The sample of 'warm color and simple' was the worst popular, so that it should not be used.

Thus, as for the direction of package design, we have 6 possibilities.

- 1) If we would like to create a good impression on the public, we should make simple but luxury packages like the 'decoration' type.
 - 2) If we want to appeal to housekeepers, we should emphasize the literal explanation in the packages.
- 3) If we value the effectiveness of package much further, we should use green color as a code of organic foods and elaborate the package.
 - 4) If we dare a novel design, we can develop the 'digital processing' type.
- 5) The classical decorative frame is still effective for a small number of purchasers who have much knowledge of the organic foods.

6) We avoid using the package of 'warm color and simple'.

6. References and Citations

- [1] Kubota, Y. (2003) Organic Foods Q&A, Iwanami Shoten, Tokyo (In Japanese).
- [2] Ministry of Agriculture, Forestry and Fisheries website about organic foods, http://www.maff.go.jp/e/jas/specific/organic.html
- [3] Keino, S. (2001) Differentiation of organic foods and its channels of distribution: a certain JA in Okayama prefecture as an example, Report of Faculty of Horticulture, Chiba University 55, pp.89-100 (In Japanese).
- [4] Collegium of Japanese Organic Agriculture (2009) A report of consumers' attitude survey on organic food (In Japanese).
- [5] Ito, S., Kiritani, Y., Ohara, Y., Tmagaki, Y., and Miyazaki, M. (2007) *Current design tendency of Japanese Sake's label*, Bulletin of Japanese Society for Science of Design 53 (6), pp. 21-26 (In Japanese with English abstract).

6. Appendix (8 clusters of 44 packages of organic soy sauce)

