Findings of Energy Label Design Research in China

Paul Waide, International Energy Agency
Christine Egan, Collaborative Labeling and Appliance Standards Program
Jin Minghong, China National Institute for Standardization

ABSTRACT

The China National Institute for Standardization (CNIS) is the agency responsible for developing a mandatory, consumer information label that designates the energy consumption of refrigerators. This label will be implemented across China and provide comparative, point-of-sale information to Chinese citizens on the energy efficiency of particular models of refrigerators relative to similar products.

In this paper, findings of a comprehensive evaluation of the proposed information labeling program in China will be presented. Our findings are based on multi-method primary research, including a consumer intercept survey, consumer focus groups, and semi-structured interviews with consumers, retailers, manufacturers, and policymakers. This paper is the first comprehensive presentation of the results of this multi-year, US $300,000 plus effort—one of the most ambitious completed to date anywhere in the world.

The focus of the project is on an energy information label for application to refrigerators, but it is expressly intended to identify elements that are important to the success of labeling for a wide variety of consumer goods, including room air-conditioners, water heaters, and motors. More than 25 energy label concepts were explored throughout the project life of just over 18 months. The goal of the research was to determine the optimal label design to achieve the maximum market impact (i.e., increasing sales of energy saving products). A further objective is the development of a label design that is: (1) easy for Chinese citizens to understand and (2) will result in improved consumer decision-making with regard to energy efficiency.

Introduction

On November 1, 1997, the National People's Congress of China approved the Law on Energy Conservation that included a mandate to develop energy information labels applicable to energy-using equipment. In response to this demand, CNIS, with technical support from PW Consulting from the UK and the Collaborative Labeling and Appliance Standards Program (CLASP) and with financial support from the Energy Foundation (USA), developed a program of work to design and implement an energy information label for refrigerators. Additional financial support was received from the Chinese Environmental Protection Agency (SEPA), United Nations Development Program (UNDP) and from the United Nations Foundation (UNF), who were involved in the implementation of a US$9 million Global Environment Facility project to encourage the market transformation of the Chinese refrigerator market. Assembling this coalition of parties and resources took time such that research on label design research began in 2002 and was completed in January 2003.

Methodology

Policymakers are generally concerned that energy labels should mitigate the informational barriers that prevent consumers from taking energy sufficiently into account when purchasing an appliance. Consumers are the primary users of energy labels and so it is appropriate that labels should be designed to present information to them in as useful and accessible a manner as possible. Clearly the format of an energy label is important in communicating this information effectively; however, it is difficult for policymakers to know what format will be most effective without research. Furthermore it can't be assumed that a label design which has been effective in one region and culture will necessarily be effective elsewhere, so successful labels are not necessarily transposable. In addition, if energy labels are to be effective market transformation instruments, they should also be sensitive to the needs of manufacturers and retailers who are responsible for the market offer. Generally, if a label is effective with consumers it will also be influential among suppliers; however, sometimes one of the most effective means of establishing how consumers are likely to respond to a design is to use the experience of those who have been supplying products to them. Lastly, the design of label needs to take into account the goals and concerns of policymakers who may wish to stress particular design elements in order to reflect national policy goals. Accordingly, the label design process should be based on research regarding the most effective design among the key stakeholders: consumers, manufacturers, retailers and policymakers.

Pre-Research Design Decisions

The project was initiated with a stakeholder workshop to discuss international experience regarding energy label design and to establish the basic principles to be followed researching the design of a Chinese information energy label. A sample of energy efficiency labels from more than 30 countries were collected and analyzed. Based on the consideration of previous label-design research findings and impact evaluation results from many countries, there was a strong consensus among stakeholders that all candidate designs should be:

1. "comparative," using a scale to indicate the efficiency of the appliance relative to that of similar actual or potential appliances providing the same service; and
2. "categorical," presenting the energy efficiency of the appliance on a scale that is divided into a number of discreet categories. This is in contrast to continuous labels that present the efficiency (or more usually the energy consumption) of the appliance relative to a single scale that is not broken into discrete categories.

International research and evaluations show that categorical labels are likely to have greater market transformation impact than continuous labels for the following reasons, (Weenig & Maarleveld 1993, Egan 2000)[A1]:

- Use of categories enables the efficiency of a product to be determined at a glance as only its *category* has to be noticed (and recalled) rather than its relative position on a scale;
- It is comparatively simple to remember the efficiency of a product during the shopping process and hence the information is more likely to be used in the final purchase decision;

- Categorical labels have a set of explicit efficiency thresholds that product designers, manufacturers, distributors and retailers can aspire to attain; and
- Promotional and marketing efforts can be targeted at specific high efficiency categories.

Developing trial labels and graphic design. In order to maximize the quality of the designs tested, four local, well-known graphic design companies were invited to develop energy label designs. These companies had experience designing logos and communications for the Chinese market and were very experienced with the cultural responses of local consumers. To ensure that they were also aware of the basic needs and concepts required of an energy label, each company was presented with a set of international energy labels for reference; they were then asked to produce their own unique designs.

This led to more than 30 trial Chinese label designs including some Chinese language replicas of leading international categorical energy labels that were to be used as benchmarks, as well as, labels that were radically different from the existing international labels. However, it was interesting that all the labels used either Arabic numbers, European letters or stars to denote the energy efficiency of the appliance concerned apparently because there is no simple way of denoting efficiency grades using individual Chinese characters. Thus, it is commonplace in modern Chinese commercial literature to use international numbers and letters to communicate rankings. It was also interesting to note that many of the unique label drafts used a mixture of Chinese characters with English text where the latter had positive associations of international acceptance and/or global thinking.

It is not practical to test 30 different label designs and so a screening process reduced this to a more manageable set of 15. Care was taken to ensure that each major design concept was represented in this subset. The initial designs are shown in Figure 1.

Iterative exploration and elimination of design concepts through qualitative and quantitative methods. The research was designed to be iterative with the dual and contrasting aims of: (1) allowing the most design concepts to be explored at each stage, and (2) progressively narrowing down the sets of viable design concepts by successive exclusion of the least successful concepts. A multi-method design was constructed to elicit feedback from consumers, policymakers, manufactures and retailers depending upon the specific task objectives as illustrated in Figure 2.

First Focus Groups

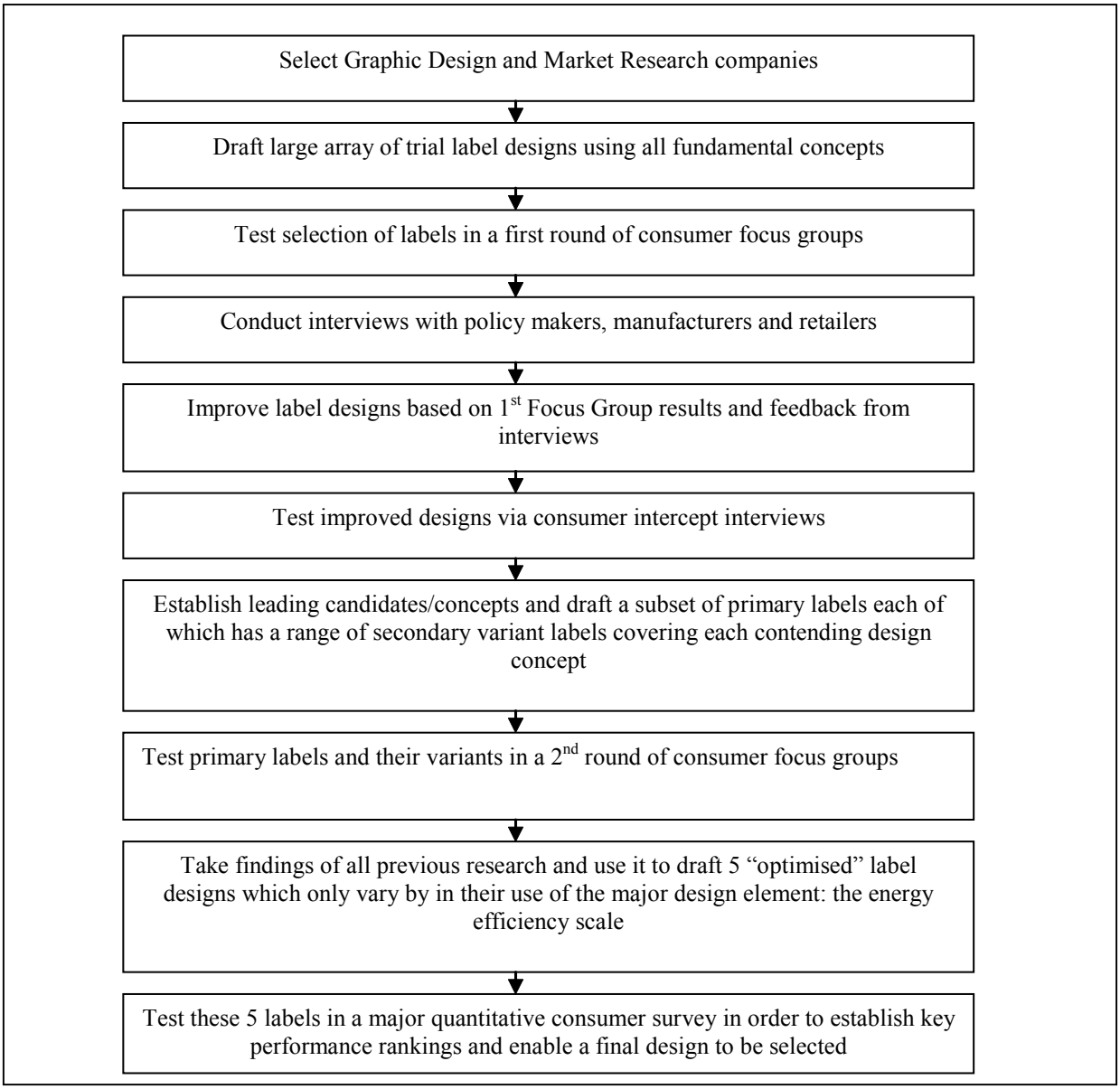
Preparation and Conduct

The objective of the first round of consumer focus groups was to establish "broad-brush" and directional feedback on initial labeling concepts. The goal was not so much to rank each of the initial candidate label designs, but to establish in a qualitative manner which elements of each label were likely to be successful and why. This type of analysis could only be done by probing responses and engaging in a dialogue. By design, the results of qualitative research such as focus groups are inherently less quantifiable than those produced by survey research (used in the later stages of the research). Thus, a number of steps were taken to ensure the quality and comparability of this qualitative work. Focus groups for two socio-economic groupings were conducted in the three major Chinese cities (Beijing, Shanghai and Guangzhou).

Figure 1. Label Designs Used in the First Round of Research

At the beginning of each focus group, participants sat around a table so that the moderator could engage them in a light conversation about an unrelated subject and thereby encourage each one to express an opinion. Once everyone had become sufficiently communicative the moderator guided them to the topic at hand by inviting discussion of factors they thought important when purchasing a refrigerator. Participants listed the main characteristics (e.g., price, brand, dimensions, freezing capacity, volume etc.) and often someone would spontaneously mention energy consumption. The moderator would then present the question of how the participants would know the energy consumption of the appliance. After considering the practicalities of ideas such as asking the retailer or looking in a catalogue sometimes one of the participants would suggest that a label indicating the energy consumption could be applied to the appliance.

Figure 2. Label Design Research Flowchart



The moderator then presented another question to the group regarding what such a label should look like and what information it should contain. In the absence of an unaided mention of energy efficiency, the moderator would enquire about the relative importance of the energy feature, the means of obtaining information about energy use, the potential value of an energy label as well as what information it should contain.

Following this discussion, meant to gradually ready participants for a discussion of energy labeling through a process called "laddering," the moderator passed out coloring pens and paper and asked each participant to design a label that could be used to communicate the energy performance of a refrigerator. This process, known as an ideation exercise, produces completely unprompted responses about what elements participants desire in an energy label.

Only after the ideation exercise was complete were participants shown pre-designed energy labels, and it was at this point that they likely became aware that the main purpose of the focus group was to help assess energy label designs. Each label design was shown in two examples side-by-side—one indicating a comparatively inefficient model and the other a comparatively efficient model. Because the sequence by which designs are shown can influence the response, care was taken to rotate the order of presentation. To begin with, the participants were shown a Chinese version of an existing international label. Specifically, the European, the Australian and the Thai label were each shown first in two of the six focus groups without mentioning that they were existing international energy labels. Only after the initial presentation were participants shown the other two international labels. Comments were solicited about each label generally along with its specific characteristics. These labels were chosen for first exposure, because they use different grading systems for the efficiency scale (e.g., letters, stars or numbers) and because they have other important design differences (e.g., dials versus stacked bars in the efficiency scale, the use of color, the choice of wording, the number of efficiency categories (7, 6 & 5), etc). Respondents were then invited to rate the features of each label from 1 to 10 in a questionnaire and to discuss them with the group.

Following exposure to each of the international labels, sets of the other labels were shown simultaneously based on thematic groupings (e.g., all the letter labels or all the dial labels). Eventually, all 15 labels were seen and evaluated by each individual as well as discussed by the group. The group was invited to comment on each design element and to suggest their strengths and weaknesses as well as potential improvements when appropriate.

Summary of Findings for First Focus Groups

- Many respondents said they would judge the efficiency just by looking at the pattern of the efficiency scale. Thus, comprehension of the energy efficiency scale should be high at first glance (i.e. it shouldn't be necessary for consumers to need to spend a long time reading the text on the label to understand the direction of the scale and whether or not the appliance they are looking at is efficient or not).
- The respondents appeared to have a very strong sense of the meaning of color such that red was clearly associated with high energy consumption, green with environmentally friendly and low energy consumption, while blue was also a color with positive environment and energy-conserving connotations.
- There was a clear association for respondents that the letter A or the number 1 denotes the most energy efficient appliance (for labels using letters or numbers respectively). The latter is because being number 1 is associated with being "the best".

- Labels using letters had the highest comprehension (100% based on unprompted questionnaire results). By contrast, a large percentage of respondents misunderstood the current Australian energy label, which uses stars to denote efficiency. There appeared to be two possible explanations for this: a) because the red color used in the dial to highlight the number of stars was strongly associated with higher energy consumption, e.g. "The red color stands for electricity consumption. The more electricity consumption, the more electricity consuming" when in fact more red was associated with more stars and hence indicated that an appliance was more efficient; b) some consumers said that as number 1 was best therefore having one star should also be best, while others associated five stars as the best (from association with the ranking used for hotels). To minimize this confusion in future trials, it was decided to produce a version of the Australian label with a green dial and with a maximum rating of five stars.
- There was also some confusion over the labels that used number scales to denote efficiency (e.g. The Thai label (No. 3) and Korean label (No.4)). A large percentage of consumers were confused by label No. 3 where the most efficient was number 5, but also many were confused by label no. 4 where the most efficient was number 1, yet the direction of the scale was seemingly counterintuitive. The preference for the number 1 being "the best" was so strong there seemed to be little point in continuing with a label where 5 was the most efficient.
- Most respondents appeared not to like having more than five efficiency categories as they found additional categories confusing.
- The use of relatively large and bold characters or numbers was preferred to smaller ones, especially when they were part of the efficiency scale.
- There was a relatively negative reaction to the inclusion of operating cost information. Many respondents said they found the operating cost "too countryside," an apparent reference to this feature as lacking in modernity or sophistication. Therefore, there seemed to be little point in including this parameter in future label designs, especially when considering the wide variance in local tariffs.
- The energy consumption value (be it daily, monthly or annual) is an important parameter to most consumers and they like to see it relatively emphasized compared with other product technical information such as the compartment sizes and noise.
- The respondents expressed a strong preference for daily rather than annual energy consumption. This surprising finding appears to be the result of prior manufacturer promotions, which had apparently conveyed the notion that an efficient refrigerator consumes less than 1 kWh per day with great success.

Manufacturer Interviews

Representatives of 20 manufacturers were contacted and asked to complete a questionnaire and interview. The respondents were generally in favor of energy labeling, especially a compulsory scheme, as they felt it would provide a level basis for competition. Their thoughts regarding the label design are summarized as follows:

- The Chinese version of the European label and the other colored bar design (Label 5 above) were strongly favored because they were deemed to be "international", clear and professional looking.

- The respondent's least favorite designs were those using stars and designs having a dial shape which were deemed unclear.
- Most respondents thought that using the national flag might confuse consumers into thinking the product was made in China rather than indicating that the label was endorsed by the Chinese state, as was the intention.

Retailer Interviews

Sixty-five refrigerator retailers from four cities were interviewed and asked to complete the questionnaire. The results largely confirmed the findings of the manufacturer interviews. Specifically, it was found that:

- The labels using stacked, colored bars (labels 1 and 5 above) were greatly favored compared to the others while the least favored were the labels using stars, because they were thought to be confusing.
- Most retailers preferred vertically stacked bars to dials although some preferred the latter.

Consumer Interviews

Following the conduct of the focus groups and other interviews, a round of consumer intercept interviews were conducted in four cities. 128 consumers were interviewed using a semi-structured research design. The purpose of this research was to: a) understand consumer's interpretation of elements in the labels, b) establish the overall interpretation of each label, and c) to learn the cause of difficulties consumers may have had in understanding the labels. Each consumer was screened to ensure that the sample of respondents was demographically and socio-economically representative of China as a whole. Each consumer was shown six out of a total of twelve label designs. These designs were related to the first round of fifteen label designs but had been improved to take into account any clear findings from the first focus group. In each city the sample of consumers was divided into four groups and each group was shown the labels in a specific but different sequence. The intention was to minimize the impact of exposure to previous labels on the responses to future labels seen in the sequence and to enable unbiased post-processing of the results. On seeing each label, consumers were asked to answer questions such as:

- Which element of the label catches your attention and why?
- What is the sense of the information in this element?
- Is it easy to understand and why or why not?
- What is the sense of information in other labeling elements?
- Compared with other appliances of the same type, do you think the electricity consumption of the appliance indicated on this label is 'a lot', 'more than average', 'average', 'less than average', or 'a little'?

Comprehension. The first label comprehension scores were lower for all the labels than in the focus group samples, which suggested there was some "training" in the focus groups due to prior discussion of the energy labeling concept and/or the group dynamic that often means once one person understand the label he/she carries the rest along. Further, the

actual comprehension of the labels, as judged by the correct interpretation of the relative efficiency of the appliance indicated in the label, was compared with the ease of comprehension scores that consumers attributed to each label and was found to be very divergent. This demonstrates that energy labels should not be designed based on what consumers say they understand, but rather on what they are proved to understand. The broader interpretation of the results suggested that consumers were likely to say they understood labels that they found were more appealing and that often the factors influencing high appeal could be applied to design elements that are actually more comprehensible.

Stars, letters and numbers. The comprehension of the letter labels was 2.5 times greater than that of the numbers and stars labels based on the first label exposure comprehension results. This may have been partially caused by cross-correlation between other label design characteristics that influence comprehension. Despite the color improvements the Star labels were still misunderstood mainly because consumers struggled to know whether the efficiency ranking was determined by the number of stars with a shaded background or by the number without a shaded background. There was also confusion because some consumers thought that the stars indicated that the label was a quality label. Out of ten consumers who were first shown the amended "current Australian" label nobody correctly understood it; however, the comprehension of the amended "old Australian" label (which used numbers to indicate the direction of the scale) was six out of nine. It was therefore proposed that a new star label should be developed that uses numbers inside the stars, as is done for the Indian energy label, and also says "the more stars the less energy used" as opposed to "the more stars the more efficient". The cause of the miscomprehension of the numerical labels was less obvious from the quotations given. Clear conclusions could not be made.

Bars versus dials. The first label comprehension test results suggested that the bar labels are twice as comprehensible as the dials; however, this may have been partly due to other factors (e.g. because the bar labels used letters or because the dial labels included stars).

The boldness of the energy consumption value. Consumers paid a lot of attention to this value and there may been a slightly higher first label comprehension result for those labels that had it indicated in bolder, clearer fonts.

Appeal. Labels with blue backgrounds scored very highly on the appeal rankings. The use of a green to red color coding in the efficiency scale was also popular.

Key Recommendations Following Consumer Interviews

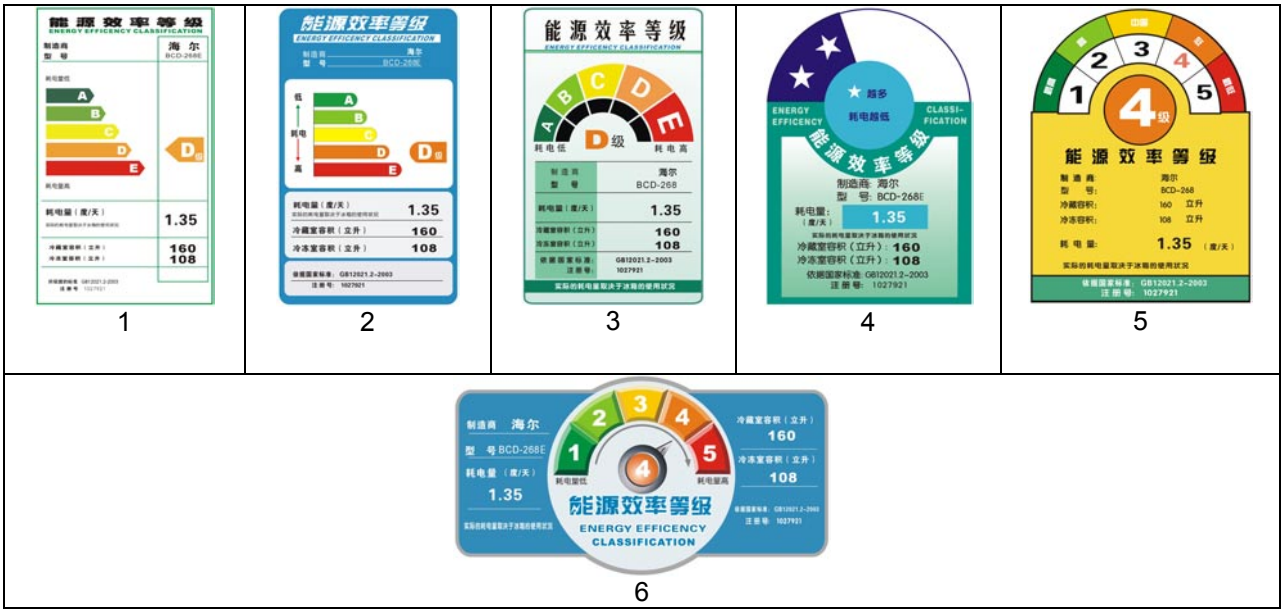
- The comprehension of some labels would benefit from having the character for 'class' or 'grade' added as is the case for the other labels.
- Each label should be tested with a green and a blue background.
- A numbered (1 to 5) vertical bar label be developed and tested.
- A star label using a dial concept should be developed where the center of the star contains a number as in the Indian label (i.e., from 1 to 5).
- The Korean label should be dropped as the amended Thai label scores fractionally higher in all cases.

Second Focus Groups

To test the several specific design elements prior to the final quantitative survey a second round of eight consumer focus groups (grouped by income, education and age) were conducted in Beijing and Shanghai. These groups were shown six "primary" labels (Figure 3) each of which had five variants making 30 labels in all. The variants addressed specific design issues such as the background color, the choice of text or the use of authority endorsement symbols. By this stage, enough information had been gathered to allow core aspects of the designs to be standardized so that all the labels had: five efficiency classes; the same title "Electricity Savings Label"; almost identical fonts and presented the same additional information. When possible, the designs used the green to red color coding, the only exception being the modified star label, which only made sense with a continuous background color to indicate the number of stars.

The results confirmed that: star labels are less likely to be understood than those using letters or numbers; labels with an Oxford Blue background had most appeal; and comprehension of scales can be aided by adding explanatory text next to each grade, or at each end of the scale.

Figure 3. Primary Labels Tested in 2nd Focus Group



1 2 3 4 5 6

Findings of the Quantitative Phase

The final round of research involved testing five optimized labels, Figure 4, via an extensive and quantitative consumer survey. Labels 1 and 2 are superficially identical but 2 uses numbers to rate efficiency in place of letters. With the exception of label 5 all labels had the same background color and all used the same fonts, information and descriptive text adjacent to the efficiency grades. Despite this, the labels vary by: numbers or letters, dials or stacked bars, vertical or horizontal layout. For the three numerical labels, number 1 always indicated the most efficient category. For dial labels the most efficient category was always at the left of the scale.

Figure 4. Final Five Label Designs Tested in the Consumer Survey

A sample of 1250 people was surveyed in urban and peri-urban locations in and around 4 major cities (Beijing, Shanghai, Guangzhou and Wuhan) and their satellite towns (Yixin, Langfan, Kunshan, Nanhan, and Xiaogan). This survey was used to select the best candidate label design(s) with statistical precision. All the consumers who took part in the research were screened to ensure they were potential refrigerator purchasers and that they would not have a potentially distorting insight into the market research due to occupation or family ties. Careful efforts were made to ensure the sample matched the demography of China as a whole in terms of sex, income, age and education. Each participant was invited to complete a questionnaire for which they received a small incentive. Labels were tested quantitatively for their comprehension and for the ease with which their efficiency ranking is remembered.

The test of comprehension involved presenting respondents with three versions of the same label execution simultaneously, each with a different energy efficiency grade (A, C and E or 1, 3 and 5, respectively), and asking them to identify which ones they would be most and least likely to recommend to a friend interested in energy use, assuming each model had similar features and operation costs were important. A strict rotation system was used to ensure that each label execution was seen 1st, 2nd, 3rd, 4th and 5th an equal number of times so that there was no bias by 'learning'. Almost 70% of respondents were able to correctly identify the most and least energy efficient model for all label executions no matter what type of person was viewing them. Despite using a large sample there is no statistically significant trend in comprehension as a function of: income, education, and age although there appears to be a strong influence according to the region (the comprehension scores in Wuhan were much worse than the rest) and to a lesser extent depending on whether the respondent has an agricultural or urban residency. The share of respondents who were able to correctly identify the most efficient model was as follows: 72% for Label 2, 70% for Label 3 and 5, 68% for Label 1 and 66% for Label 4. The share of respondents who were able to correctly identify both the most efficient and least efficient model was: 64% for Label 2, 63% for Label 3, 60% for Label 1, 58% for Label 5 and 56% for Label 4. Overall, the comprehension results of Labels 2 and 3 are significantly higher than for Label 4 while those of Label 2 are slightly higher than for Label 3. It should be stressed that these high comprehension scores are the product of the earlier qualitative and semi-quantitative design work, which had eliminated ineffective design concepts prior to the quantitative analysis stage. As roughly 70% of people can understand these labels without having seen them before, a very high comprehension rate could be expected following a promotional campaign and after a single label design has been in place for a number of years.

Perceptions of Label Appeal, Motivation, Credibility and Comprehension

Consumers were asked to rank on a 1 to 10 scale their ratings of each label execution for its ability to capture attention, its credibility, the appropriateness of the level of information, its ability to motivate the user to consider energy efficiency when making a purchase and its ease of comprehension. The questions were posed following exposure to individual label executions and reposed after exposure to all the labels. The individual exposures produced quite similar scores for all the labels with no label producing an average score of worse than 6.71 and none higher than 8.03. Again, these results suggested that all of the five final label designs worked well and indicated that lessons had been learned through the earlier design sessions. By contrast, when all five labels were shown simultaneously, Label 4 scored far higher than the others with 35% of respondents choosing it as the one they would be most likely to read and 38% the most motivating, while only 15 to 16% give it the lowest rating for these parameters. It is very revealing that a similarly high share of respondents (35%) thought it was the easiest label to understand even though the true comprehension tests reported above found it was the least likely to be correctly understood.

This demonstrates an important factor in energy label design research: that consumer perceptions of which label is easiest to understand do not necessarily correlate with their actual levels of comprehension. In this case, it is quite possible that many of the factors they found appealing about the design were actually distracting them from the main message of the label.

Memory Tests

The test of memory involved presenting respondents with five versions of the same label execution, each with a different energy efficiency grade and other lesser variations for the model name, manufacturer and daily energy consumption, in order to test the respondent's ability to correctly remember all of the five efficiency grades. The analysis showed there was no significant difference in 'recall' for the different label executions with all scores being between 26 and 30%.

Conclusions

Although Chinese consumers were initially unfamiliar with energy efficiency labeling, once they were exposed to the idea they clearly support it and pay great interest to the information a label presents. To summarize, the final Labels 2 and 4, each had their strong points but Label 2 surpasses the other four labels in terms of its ability to be correctly understood, which is the most important criterion. The fact that some 72% of respondents who had never seen the label or other energy labels before were able to correctly identify that an appliance had a relatively high efficiency demonstrates the value of structured label design research in developing an effective communication vehicle. Overall, the project has highlighted the need to base energy label designs on market research and has demonstrated the need to base research results on more than apparent label popularity.

The Chinese information energy label will initially be implemented voluntarily for refrigerators from 2004, before becoming mandatory in due course. After refrigerators the plan and intention is to progressively extend application of the energy efficiency label to other products with substantial energy saving potential.

The authors would like to thank all the people who contributed to this project and the various agencies that supported it.

References

W. Weenig & M. Maarleveld *Impact of Energy Information Labels on Consumer's Choices of Refrigerator*s, E&M Centre for Energy & Environmental Research, Faculty of Social and Behavioural Sciences, Leiden University, E&M/R-93/36, The Netherlands.

C. Egan. "Testing of International Appliance Labeling Approaches with U.S. Consumers." Paper Presented at the Second International Conference on Household Appliances and Lighting (EEDAL 2000), Naples Italy.