

An Evaluation of the EnergyGuide Label: What we learned.

*Sandy Smith and Jennifer Thorne
American Council for an Energy-Efficient Economy*

ABSTRACT

This paper critically overviews a project designed to evaluate the EnergyGuide Appliance label that was launched in 1980 by the U.S. Federal Trade Commission. In particular this research sought to measure the efficacy of the existing EnergyGuide label, develop a range of alternative labels, and conduct comparative studies of the existing and alternative labels.

Seven interrelated studies were conducted to respond to these questions. The form of these studies varied from semi-structured interviews and focus groups to a survey and quasi-experiment. Participants included supply-side actors such as white goods manufacturers, HVAC contractors and sales staff, and consumers of white goods. The research was conducted in 13 cities across the U.S. In this paper we will describe briefly each of the studies, identify the respective benefits and limitations, and critique both the process that was undertaken as well as the findings in relation to other research.

The project identified a number of features of product labels that are preferred by participants. In particular, preferences for categorical and continuous scales were measured. Of the categorical scales, the stars-based rating emerged as the clear preference, especially as the stars were congruent with the notion of an ENERGY STAR certification. Of the continuous designs, a bar graph with scale marks was preferred, although it was not as popular as the stars-based categorical measure.

With regard to the informational elements presented on the label, annual operating cost was considered the most important piece of information, although annual energy use was also considered important. The visual clarity and placement of this information was also found to be salient. Participants were critical of the amount of explanatory text on the original label, although there was no consensus as to what information was unnecessary. Minor editing of the content met with greater approval.

Overall, despite a reasonable familiarity with the EnergyGuide, it appears to have limited impact on product choices of US consumers. Findings provide strong evidence that the EnergyGuide can be redesigned to improve consumer comprehension, encourage wider use of the label, and motivate consumers to consider energy use when purchasing a labeled appliance.

Introduction

In the United States, the Energy Policy and Conservation Act of 1975 and the National Energy Conservation Policy Act of 1979 directed the U.S. Federal Trade Commission (FTC) to develop a labeling program for certain home appliances and energy-using equipment. The program was launched in 1980 with two legislated goals: to improve energy efficiency and assist consumers in making purchase decisions. In fact, the ability or inability to assist consumers is listed as a primary criterion for deciding to label specific products. The current EnergyGuide label is displayed in Figure 1.

Through small-scale studies and anecdotal evidence over the last twenty years, concern has risen that the U.S. label may not be living up to this legislative mandate. No systematic evaluation of the EnergyGuide program or the current label design has been undertaken. In addition, over the last five

years, alternative approaches to appliance labeling have been developed and implemented elsewhere in the world with impressive results in terms of consumer awareness, market impacts, and energy savings.

In this context, ACEEE, with input from other organizations, decided to evaluate the efficacy of the EnergyGuide label and determine the best label format and graphical element for U.S. consumers. Primary research with consumers sought to determine the best label format and informational elements for U.S. consumers. In addition, supply-side actors (e.g., manufacturers, contractors, and retail sales staff) were interviewed to uncover opinions regarding program efficacy and the optimal label format. This paper reports the research that was undertaken to achieve these aims.

Method and Results

A multi-method, sequential research design was constructed to elicit consumer feedback as summarized below. Seven studies were conducted in total, the first with supply side actors, and the remainder with consumers. The consumer studies culminated in a quasi-real world shopping experience. Interviews with supply-side actors and consumers (studies 1 and 3, respectively) were conducted by ACEEE staff and contractors. The remainder of the qualitative research and all quantitative research was conducted by an independent market research firm. This paper will discuss only those findings directly related to the label design, for more complete coverage of the results and methodology, please refer to Thorne and Egan (2002). The qualitative studies were designed to gather directional feedback on the current label and alternate designs with an emphasis on label preferences, opinions on various informational elements, and comprehension and interpretation of the labels. The most popular themes and preferences emerging from these research tasks are reported. For the quantitative studies, statistical tests conducted with an alpha level of .05 are reported as significant unless otherwise noted.

Study 1: Interviews with Supply-side Actors

Manufacturer Interviews

Methodology. ACEEE conducted 16 telephone interviews with representatives of white good appliance and heating, air conditioning, and water heater equipment manufacturers in order to document: (1) how manufacturers perceive the current EnergyGuide labeling program; (2) their thoughts on whether and/or how that label is or isn't working; and (3) their experiences in implementing the EnergyGuide label. Participants were not shown the current label or alternatives as part of their interview.

Summary of key findings. Overall, the interviewees were highly experienced with regard to the label, often with a long history of direct work in the program. Nearly unanimously, respondents felt strongly that the program was not achieving its intended informational goals and they challenged the usefulness of a label for products where the appliance and label typically are not seen at the point of sale (i.e., installed appliances such as heating and cooling equipment.)

The majority of respondents felt the label had enjoyed some, though not complete, success. However, nearly all the interviewees felt that the program's effect on the market is minimal. Respondents enumerated several strengths of the EnergyGuide label, including the label's simplicity (often linked with ease of understanding), accuracy, prominence, consistency from product to product, and provision of comparative information. Both white good and HVAC manufacturers commented negatively regarding the label's message. Several respondents indicated that the label was not consumer-friendly. The label was seen as a poor communication tool because it was either overly technical or

unattractive. Technical accuracy was another significant area of discussion. In particular, the interviewees challenged the way products are grouped in the categories that form the basis for the comparisons on each label, and the use of national average figures.

HVAC Contractor Interviews

Methodology. ACEEE conducted nine telephone interviews with HVAC contractors to assess: (1) the role of the EnergyGuide label and energy efficiency in the purchase of installed appliances (i.e., central air conditioners, heat pumps, furnaces, boilers, and water heaters); (2) how contractors perceive the current EnergyGuide label; and (3) their thoughts on whether and/or how that label is or isn't working. Participants were not shown the current label or alternatives as part of their interview.

Summary of key findings. According to most of the interviewees, energy efficiency is a factor in customers' purchase decisions; however, there was wide variation in the relative importance the interviewees felt customers actually placed on efficiency versus other features. Most importantly, however, none of the interviewees used the label as a source of energy efficiency data and only one reported actually using the label as an informational tool with customers when selling HVAC equipment. By contrast, all of the respondents indicated that they provided customers with efficiency information from the manufacturers (e.g., brochures). The majority of respondents felt the label was of no use to their customers because the information is presented after the purchase decision is made (i.e., when the equipment is installed). Surprisingly, although contractor interviewees were overwhelmingly negative regarding the label's efficacy, they were optimistic regarding the potential of a revised labeling or other information program.

Retail Sales Staff Interviews

Methodology. ACEEE conducted 16 in-person interviews with sales staff of a major retail chain to assess: (1) the importance of energy consumption as a factor in the purchase of labeled appliances; (2) interpretative capabilities of the information presented in one of five appliance labels (four alternatives plus the current); and (3) the preference of label format among the five choices.

Summary of key findings. Energy consumption was mentioned as an issue in appliance sales in many of the interviews. The interviewees felt that concern for energy efficiency was relevant to a subset of customers, particularly those shopping for refrigerators and room air conditioners. The interviewees indicated that they particularly used the label in situations where energy efficiency was a deciding variable between two otherwise equivalent products.

The retail sales staff was the only group of supply side actors that were shown the current EnergyGuide and the alternate labels for interpretation and side-by-side comparison of preferences.¹ The majority of the salespeople interviewed preferred a categorical label as the best display of the five alternatives shown to them. These respondents thought that the categorical style would allow consumers to interpret the label more quickly and easily. However it should also be noted that participants were drawn from the staff of a major chain that includes training on the EnergyGuide as part of new staff training.

In summary, the retail sales staff indicated that they made use of the current label in at least some of their sales, particularly with customers interested in energy efficiency and in cases where energy use

¹Retail sales staff were shown the same labels as the consumers interviewed in Study 3.

helped to differentiate among products. The interviewees were very open to further improvements to the label and a majority felt that a categorical style would be easier than the current option for consumers.

Study 2: Consumer Focus Groups- Round 1

Methodology. ACEEE contracted an independent market research firm to conduct an initial round of six consumer focus groups (four with white goods shoppers and two with larger household equipment shoppers) to examine consumer perceptions of the current EnergyGuide label and responses to alternative label designs. Pairs of labels were shown side-by-side, one representing a case of low energy use and one a high. The labels used in the focus groups were left very close to their original format. The groups were not expected to come up with final designs, rather they were intended to uncover trends and general directions for additional qualitative and quantitative analysis.

Summary of key findings. The groups found that the current label, though familiar to the participants, was not always read or used in making appliance purchase decisions and improvements were suggested. Overall, the groups suggested that an ideal EnergyGuide label would:

- Include and highlight the estimated annual operating cost so it can be easily seen;
- Include and highlight the annual kilowatt-hour (kWh) usage so it can be easily seen;
- Use a yellow background as this is recognized and associated with energy information;
- Use a visually appealing graphic that simply and clearly communicates the kWh usage;
- Include appliance specifications such as the manufacturer and model number;
- Reduce the amount of unnecessary text;
- Clearly state that the label is regulated by the U.S. government; and
- Be formatted and outlined to communicate its messages using blocked-off spaces and relationally grouped information.

These suggestions were used to improve each of the designs for later tasks. In addition to these graph-specific questions, the groups explored the addition of the ENERGY STAR to the EnergyGuide label for qualifying products. In these groups, the low recognition of ENERGY STAR and its purpose was problematic. Thus, it appeared that until recognition is higher, the addition of the ENERGY STAR might be of limited use in encouraging energy efficiency.²

In summary, most of the participants reported a relatively low priority on energy efficiency in appliance sales and a low level of use of the current EnergyGuide label. However, they also indicated that problems with the current label limit its usefulness and appeal (i.e., it is too cluttered, poorly organized, and overly technical). The participants made various suggestions to improve the current label as well as the other alternatives. In general, these improvements were geared toward making the labels simpler and more direct in communicating their main message as well as more graphically appealing.

Study 3: Consumer Interviews

Methodology. ACEEE completed a total of 54 semi-structured customer intercept interviews in three cities: Boston (28 interviews), Denver (18 interviews), and Dallas (8 interviews). The interviews in Boston and Denver were with customers shopping for white good appliances (refrigerators, freezers, dishwashers, clothes washers, or room air conditioners), while the interviews in Dallas were with customers shopping for water heaters. In each interview, four improved label designs were tested (based on the results of the first set of focus groups), in addition to the current EnergyGuide label. The purpose

² Since the time this study was conducted (June 1999), market research has shown a substantial increase in recognition of the ENERGY STAR label.

of this set of consumer interviews was to evaluate the current label in-depth and side-by-side with alternative labeling approaches to draw out comprehension and information-processing issues as well as to examine the reasons behind reported label preferences. The order of presentation was systematically varied to control for order effects.

Summary of key findings. The majority of the interviewees (roughly 75 percent) correctly interpreted the single graph that was presented to them first as a test of comprehension (i.e., they deduced that the model depicted was a high energy-using model). The current label had the highest rate of misunderstanding while the speedometer had the lowest. The star, thermometer, and letters graphs fell in the middle. Also, while the majority of participants were able to deduce that the model depicted was not very energy efficient, fewer could articulate or use the graph's comparative element. In addition, a few of the respondents mistook the operating cost figure for a savings number. However, this previously identified comprehension problem (du Pont 1998) was less frequent than expected, perhaps due to modifications made to all designs except the current label to more clearly identify the operating cost.

No clear winners emerged from the analysis of the interviewees' preference-related statements. Another area of overarching comment was in the expression of the basis for the comparison. Several participants suggested that at a minimum all of the labels should indicate clearly that the comparison was based upon a range of similar models. However, many suggestions were given for improving all types of labels

In summary, although all the label formats were comprehensible to a majority of the interviewees, the current label appeared to be the most difficult for the interviewees to interpret. Furthermore, the interviews support the conclusion that at least from a consumer perspective, improvements over the current label are possible. This is evident in the relatively high incidence of comprehension problems with the current label and the relatively low incidence of preferences for the current label over all of the other options. The thermometer and speedometer were promising label options due to their graphics although the similarity in interviewees' perceptions of these two designs suggested that they were not different enough to warrant continuing to test them both. The stars label appeared promising because of its intuitive scale as well as its strong motivational potential. The letters label appeared to be the least refined of all the graph designs and needed improvement and continued testing. An overall comprehension problem was that most people did not immediately grasp that the model in question was being compared to other similar models.

Study 4: Consumer Focus Groups-Round Two

Methodology. ACEEE contracted a second round of six focus groups to examine multiple executions of each of the leading label designs. The focus groups consisted of owners of single-family homes in the market for household appliances and equipment sold through retail stores. The findings from these groups were expected to lead to final designs for use in quantitative testing. Four basic graph alternatives were tested including the current label, the star-based label, the thermometer-based label, and the letter-based label. These labels had been modified to reflect many of the improvements and suggestions drawn from the earlier consumer tasks.³ In addition, variations on these basic graphs were tested that incorporated additional informational or visual elements.

³ The speedometer was dropped due to its poor testing in the initial round of focus groups and, based on the consumer interviews, its seeming overlap with the thermometer. Also, manufacturers who had seen this version as part of a Canadian study indicated that it would be very difficult and expensive to implement.

Summary of key findings. Respondents viewed the stars label most favorably. The star graphic was considered consumer-friendly because it was simple to interpret and most consumers were already familiar with the concept of using stars to rate products and services. Many respondents noted that the star graphic easily and effectively communicated the energy efficiency concept to consumers. However, although the majority of consumers found the star graphic highly effective at communicating the intended message, many noted that the basic version was not very informative. Thus, most group members preferred executions that increased the amount of information available on the label.

Participants indicated that they liked the level of information contained in the current EnergyGuide in spite of the fact that the graphic is relatively ineffective. Furthermore, they indicated that they were familiar with the current label and believe it is easily recognized by consumers. Generally, respondents found it difficult to understand the inverse nature of the relationship between kilowatt use and efficiency. Thus, while they preferred that energy use be reported in terms of kilowatt-hour usage rather than EER, they also want the label to reflect better performance (i.e., most efficient) at the top or right of the scale and lower performance (i.e., least efficient) at the bottom or left of the scale. This finding suggests that it is difficult to satisfy consumers' demands for a scale measure that they are both familiar with and find intuitive.

In summary, the participants preferred stars over the other graphical options with the current label being the second most preferred. The stars were strong visually and from an information-processing perspective were very clear and quick. The current label was strong because of its familiarity and depth of information. Overall, the participants wanted a label that incorporated both a strong graphic and detailed information so designs that combined these features were most preferred.

Study 5: Consumer Focus Groups-Round Three

Methodology. ACEEE contracted for a third round of four focus groups to evaluate additional continuous and categorical rating concepts and determine the preferred level of content for the EnergyGuide label. The focus groups also explored how categorical rating systems interact with the ENERGY STAR given a recent ruling by FTC that allows manufacturers to add the ENERGY STAR to the EnergyGuide label. In addition, the groups investigated consumer preferences for where to place the ENERGY STAR on the EnergyGuide label. The groups were intended to provide input on the design of the labels to be tested during the following quantitative research phase consisting of a consumer survey and simulated shopping experiment.

As in the earlier tasks, the label executions were modified to address issues and suggestions raised in previous research tasks. The thermometer and European-style letters label were dropped in response to the negative responses in the second round of focus groups. In addition, new variations on the label were included to test additional graphical rating concepts. Four continuous-style labels were tested including the current (or line graph) label, the bar graph, the gradation graph, and the slope graph. Three categorical-style labels were tested as well, including stars, check marks, and letters. The current label was tested with the level of explanatory text that is currently found on the actual label. All other designs were tested with a medium-high level of text based on input from the earlier focus groups. Versions of the current, stars, and check marks labels were tested with the ENERGY STAR. Unlike the earlier tasks, the current label was not identified to participants as the EnergyGuide design in current use; it was referred to as the line graph label.

Summary of key findings. Respondents found the stars label the most effective rating concept, preferring it over all other designs. Like the other categorical rating concepts, respondents found the stars label easier to read and understand than the continuous graphing concepts. Specifically, the stars label was considered most attention-grabbing and most effective at increasing consumer awareness of

energy efficiency. Participants appreciated the familiarity of the stars rating concept, noting that the stars would allow consumers to quickly judge which models were more efficient without reading the fine print. Furthermore, respondents felt a star rating would be most likely to encourage consumers to purchase a more efficient appliance. While many group members recognized that the stars rating was strictly limited to the efficiency of the appliances, others noted that the use of stars might infer a rating of quality and performance beyond energy efficiency.

Despite its similarity to the stars label, the majority of participants did not respond favorably to the check marks label, which they found visually unappealing. Some respondents appreciated the novelty of check marks as opposed to the common stars rating; however, most felt the check marks did not carry the same weight as the stars in communicating the importance of energy efficiency or influencing purchase decisions. Similarly, reactions to the letters label were generally negative.

Among the continuous-style labels, the bar graph was favored over the alternative graphs by virtually all respondents. The gradation and slope graphs were dismissed almost immediately because respondents found them unattractive, confusing, and subject to misinterpretation. Respondents also considered the bar graph much more user-friendly than the current line graph. The strongly stated preference for the bar graph stems from the use of a well-defined arrow as opposed to a carat-style pointer to signify the labeled model's energy use, use of the bar which varies in length according to energy consumption, and the inclusion of tick marks along the full range of the scale. Relative to the current label, the bar graph was considered more visually appealing and easier to understand but was not found to be more attention-grabbing or more effective in encouraging energy-efficient purchases.

Respondents voiced other concerns about the current line graph label. Many did not recognize the line graph on the current label as a graphical element and were unable to use the graph to make comparisons to the energy use of similar models. Some consumers mistakenly identified the operating cost information on the label as an estimate of energy savings. The high level of text on the current label appeared to contribute to its lack of visual appeal and discouraged many participants from reading the label. A slight majority of respondents preferred the medium-high level of text to the high level and believed this provided the level of information needed in appliance purchasing.

As mentioned above, for the first time in our research, the current label was not identified to respondents as such, but was merely referred to as the line graph label. Few participants recognized the label or recalled seeing it when shopping for appliances. Overall, respondents reported that they had not read the label or used it as an information tool when they were in the market for appliances.

Few of the respondents were familiar with the ENERGY STAR. Despite this lack of awareness, participants interpreted the label as an endorsement rating for appliances that are above average for energy efficiency. Respondents reacted negatively to any placement of the ENERGY STAR within the box containing the line graph on the current label. Such placement was confusing to respondents because it was not clear whether the ENERGY STAR was an indicator or device corresponding to a point on the line graph. Furthermore, participants indicated that the label cluttered the graph, making it more difficult to read and interpret. Placement of the ENERGY STAR in the bottom right-hand corner was vastly preferred and many respondents indicated that it would be more effective in this location. Respondents also believed that the ENERGY STAR and the various categorical rating schemes' stars, check marks, or letters complemented each other and worked well together.

In summary, the stars label was the preferred categorical label and the preferred label overall while the bar graph outperformed the other continuous-style labels. The stars label was visually appealing, attention-grabbing, and easy to comprehend quickly and from a distance. In addition, the stars label effectively demonstrated the concept of comparison of energy use among models. The bar graph was more user-friendly and improved comprehension relative to the other graphing concepts.

Respondents preferred a medium-high level of text, which they found to provide the appropriate level of information without cluttering the label and discouraging its use. The ENERGY STAR was

understood to be an endorsement label and considered a complement to the EnergyGuide rating system, particularly when placed in the bottom right-hand corner of the label.

Study 6: Consumer Survey

Methodology. ACEEE's contractor completed a total of 500 surveys with homeowners through mall intercept interviews in five cities to determine which among the lead label concepts has the highest rate of comprehension and motivating ability (in terms of encouraging consumers to read the label and consider energy efficiency in their purchase). Based on the results of the third round of focus groups, five label designs were tested, including three categorical designs (letters, stars and checks) and two continuous designs (the current line graph label and a modified bar graph including scale marks).

The first comprehension questions involved showing the respondents a single label design (with the model and energy performance held constant for all five designs) and asking them to identify the level of energy efficiency depicted. The next test of comprehension involved presenting respondents with three versions of the same label execution, each with a different kWh level, and asking them to identify which one they would be most likely and least likely to recommend to a friend, assuming each model had similar features and operating costs were important. Under this scenario, which is considered to be the easiest test, three versions of the same label were presented side by side. Finally, in the third comprehension test, respondents viewed five label executions, one at a time, and reported the impact of the label on a purchase decision. They rated each of the five labels, one at a time (order was rotated), on ease of understanding using a 10-point rating scale (where 1 signified 'not at all easy to understand' and 10 signified 'extremely easy to understand').

Summary of Key Findings. For the first test, the stars and the checks labels had the highest rates of correctly communicating the energy efficiency level of the appliance they described, while the current label and the letters label had the lowest rates. The bars label had a moderate accuracy rate compared to these other types of labels. These results were found to be statistically significant with an alpha of .05 after conducting an Analysis of Variance (ANOVA) on the response rates.

For the second test, more than three-quarters of respondents were able to correctly identify the most and least energy-efficient model for all label executions. However, respondents were more able to correctly identify the most efficient model using the checks label than the letter label and also were more likely to give no correct answers with the bar graph label than the stars label. Furthermore, when an ANOVA was run to determine whether the number of correct answers is the same for all label executions, significant differences were found across the label types. Thus even though the differences appear to be minor, they were supported by the statistical tests.

For the third test, an ANOVA was used to compare the mean comprehension ratings of all five labels, and a significant main effect was found, indicating that there were significant differences in the ratings across the labels. Paired comparisons found that the stars and checks labels were easier to understand than all other labels. Furthermore, the letters and the bar graph labels were also rated as being significantly easier to understand than the current label.

After rating each label one at a time, all five label executions were placed before respondents at the once in the order (left to right) that the respondent originally saw the five labels. Then respondents were asked to choose which one label would most motivate them to consider energy use in their appliance purchase. The stars label was judged the most motivating (45%) followed by checks (20%). The bar graph (14%) and letters (12%) were less motivating, and the current label (4%) was least motivating by a wide margin. Thus, the survey suggested that the stars label was the most motivating to respondents followed by checks and the bar graph. Taken together, these results show that the stars and the checks label were, in fact and perception, rated as most easily understood by respondents.

In summary, the survey results suggest that the best label design for U.S. consumers in terms of ease of understanding and motivating ability is based upon stars. Checks are a close second but appear to have less motivating capacity than stars. The bar graph was an improvement over the current label but not as easily understood or motivating as either stars or checks.

Study 7: Simulated Shopping Experience

Methodology. A total of 204 homeowners were recruited to ‘shop’ for clothes washers and water heaters in a simulated retail shopping environment. The appliances on the mock shopping floor included prices, feature cards, and manufacturer point-of-purchase materials. The participants, including 49 married couples, were asked to select the model of each appliance that they would most likely purchase and answer a series of questions about their choices. Each participant saw the stars label on one appliance type and the bar graph label on the other; the order in which each label design was seen and evaluated was varied across respondents.⁴ Respondents were not notified that energy use and the EnergyGuide label were the subjects of the study.

Summary of key findings. All statistical tests in this study were conducted with an alpha of 0.1 as a result of the likelihood of small cell sizes resulting from the research design. Overall, there were no systematic differences in the products most likely to be purchased based on the label displayed. However, the most efficient washing machine was more likely to be selected when accompanied by the stars label than the bar label. When asked the reasons for their purchase preferences, respondents most frequently cited energy efficiency as the reason for their water heater selection. For clothes washers, energy efficiency was the fourth most frequently cited reason after features, price, and capacity.

In general, respondents’ perceptions of the quality of the appliance models in the study did not differ in any systematic way by label type.⁵ However, there were a few individual differences. The second most efficient water heater was perceived to be of higher quality by respondents who saw the bar label on water heaters than by those who saw the stars label. Respondents who saw clothes washers with the stars label perceived the fourth most efficient washer (the least efficient ENERGY STAR model) to be of higher quality than those who saw the bar label. The least efficient clothes washer was rated at below average quality more often when displayed with the stars label than the bar label.

Perceptions of appliance value also revealed little variation by the label displayed. Respondents were more likely to consider the least efficient clothes washer, as well as the least efficient of the ENERGY STAR models, as a below average value when the models were displayed with the stars label than the bar label. On the other hand, a relatively inefficient clothes washer model was considered to offer higher than average value by a significantly higher portion of who saw the model with the bar label. These findings seem to suggest that the stars label helped respondents distinguish poor values among the less efficient models. There were no significant differences in value perceptions by label for the water heaters.

Energy efficiency was an important determinant of appliance quality and value regardless of the label displayed. On an unaided basis, respondents most frequently named energy efficiency as an attribute differentiating between above and below average quality water heaters. Energy efficiency was the second most frequently named attribute of quality, after features, for clothes washers. When

⁴ Since earlier studies strongly supported the need for improvements to the current EnergyGuide label, and the stars label and the bar graph label emerged from earlier tasks as the optimal categorical and continuous style designs, respectively, these labels (shown in Figures 2 and 3, respectively) were selected for this final test.

⁵ The survey instrument did not define *quality* or *value*—shoppers assigned their own meanings to terms which we treated as independent variables in our analysis.

considering appliance value, energy efficiency was the most frequently named determinant of above and below average value for both clothes washers and water heaters. The importance of energy efficiency in participants' perceptions of appliance quality and value was the same for each label.

In summary, the differences among the appliances most likely to be purchased based on the label design were modest. There is some evidence that respondents were more likely to purchase an efficient model when they saw the stars label than when the bar label was displayed. Label design had no systematic impact on respondent perceptions of appliance quality. However, the stars label did communicate less value for the inefficient models with higher operating costs.

Overall Findings

Graphical Elements

Overall, the research tested multiple types of graphical comparisons of energy use including two broad labeling approaches: categorical versus continuous scales. Stars emerged as the most preferred categorical rating element, and the most preferred label design overall. Consumers are familiar with star ratings and believe they are easiest to use and comprehend quickly from a distance. Stars were also found to be most motivating to encourage consumers to use the label and consider energy use in their appliance purchase. Other categorical rating schemes, including letters and check marks, have confusing meanings and other associations (e.g., school grades, checklists) for consumers.

The bar graph with scale marks represented the most significant improvement over the current line graph. Although respondents liked the use of continuous-scale device images (e.g., speedometer and thermometer) and often preferred them to the simple bar graph, these images raised questions about the relevance of the device to energy use. These concerns outweighed the visual appeal of the device images in favor of the bar graph.

Another notable finding is the analytically complex nature of energy efficiency for many consumers. The inverse nature of the relationship between energy efficiency and energy use makes it challenging to devise a graph that clearly and quickly explains the issue to everyone. In particular, it seemed that vertical scales such as the thermometer and the letters were difficult for consumers to intuitively understand. Interestingly, the scales on these two labels were reversed. In one, the best product was on top while in the other, the worst was on top. Yet in both cases some consumers felt that the scales should be reversed.

Informational Elements

There were some interesting contrasts in the way different stakeholders interpreted the labels and in particular the relative priority placed by the various stakeholders on specific informational elements of the label. For example, for many of the manufacturer interviewees the strength of the current label is its accuracy and ease of understanding. Many of the manufacturer interviews found annual kWh usage to be a particularly useful and accurate element of the label. These manufacturer responses on the simplicity of the label and in particular kWh as a measure of energy use are in conflict with much of what was discovered from consumers and retail sales interviews.

In the early studies, consumers reported that annual operating cost was one of the most important informational pieces on the label. They were also interested in annual energy use data. In both cases, participants suggested that operating cost and energy use information be highlighted and each clearly blocked off and labeled so it can easily be picked out from other label information. Taking this suggestion into account, all of the designs except for the current EnergyGuide were modified prior to the consumer interviews in order to more clearly identify the operating cost figure.

Level of Explanatory Text

Overall, consumers and retailers found the current label complex and/or overly technical. Furthermore, a significant proportion of consumers had difficulty understanding it. Although some consumers liked the detailed nature of the label, many indicated that they don't usually read all of the text and furthermore that they don't necessarily use the label at all. Indeed, some reported that the level of text made the label appear intimidating and discouraged them from reading it or using the information it provides. Unlike manufacturers, consumers emphasized the importance of dollars over kWh in their analysis of the labels. The importance salespeople placed on operating cost and the appeal of categorical labels as simplifying the explanation of energy use in the sale echoes the views of consumers.

In early focus groups and consumer interviews, many respondents recommended cutting the amount of text on the label to make it more visually appealing and less intimidating. However, participants also liked getting the maximum amount of information from the label. These findings highlight some of the conflicting demands that study respondents placed on the label.

In response to these findings, two levels of text were tested ("high" corresponding to the current label content and "low" with a minimal amount of text) were tested in subsequent research tasks. Respondents were asked which level of text they preferred, what specific content was essential to the label, and what text, if any, was not useful and could be eliminated. Despite the view that the label was too busy, respondents were reluctant to have information taken away. In the end, a medium-high level of text, including most of the original content, emerged as the preferred variation. This design provided the information consumers believed the label needed to convey but eliminated the one sentence that was found to be obvious and of little value, thereby slightly reducing the amount of text.

Perceptions of Product Quality

Based on comments from a few focus groups participants and survey respondents, there was some concern that a categorical rating system, particularly the stars-based rating, might mislead consumers by implying a rating of product quality in addition to energy efficiency. The simulated shopping experiment tested the impact of the stars label and the bar graph label—the optimal categorical and continuous label designs identified through earlier research tasks—on consumer purchase decisions and on perceptions of product quality and value. Overall, there were no systematic differences in the products most likely to be purchased based on the label displayed. Furthermore, label design had no systematic impact on consumer perceptions of appliance quality or value, although our findings did suggest that the stars label helped respondents distinguish poor value among the less efficient models with higher operating costs.

Interaction with Endorsement Label

Respondents easily distinguished the ENERGY STAR from the categorical rating schemes, recognizing that ENERGY STAR is an endorsement that the appliance has met prescribed standards, while the categorical rating is a scale for the comparison of energy use among models. Indeed, there appeared to be a mutually reinforcing relationship between the ENERGY STAR and the categorical and continuous labels with some improvement in comprehension of the labels with the ENERGY STAR relative to labels of the same graphical design displayed without it.

In general, respondents reacted negatively to the current placement of the ENERGY STAR within the box containing the line graph. This placement was confusing because it was not clear whether the label was an indicator corresponding to a point on the line graph. Furthermore, the ENERGY STAR cluttered the graph, making it more difficult to read and interpret. Placement of the ENERGY STAR in the bottom right corner of the EnergyGuide was vastly preferred and many respondents indicated it would be more effective in this location.

Summary of Findings

A common theme that has emerged from nearly every study is that while energy efficiency receives some attention from the various audiences, it is not typically a primary driver. In addition, although all of the market players are used to and familiar with the current label, there appears to be a low level of use and, on the supply side, perceived value. Most importantly, the current label also appears to have minimal impact on consumer, manufacturer, and contractor comparisons and choices. Indeed, when the current label was not identified as such to consumers, most were unable to identify it or correctly select it from a group of different label designs.

Among the categorical labels tested, a clear preference among consumers emerged for a star-based rating system. Most consumers found the stars rating most easy to understand and most motivating. Letters and checks often had confusing meanings to consumers, whereas the use of stars was more recognizable and intuitive. Furthermore, consumers found the stars rating system complementary with the ENERGY STAR certification. Of the continuous label designs, the bar graph label (Figure 3) was preferred to the current label design but did not test as well as the stars-based categorical label (Figure 2) for visual appeal, attention-grabbing ability, ease of understanding and use, and motivating ability. Importantly, neither of these labels was found to mislead consumers by implying quality or other characteristics beyond energy consumption.

A Critical Overview of the Process

So far, we have discussed the actual findings of the research project. However, of equal importance for further evaluation work of this type are the conclusions about the research process that was undertaken. An important feature of this project was the multiple study design that was deliberately sequential and cumulative in nature. Studies 2-7 were designed so that knowledge gained in each study fed directly into subsequent studies. This knowledge was used to fine-tune both the research questions and the stimulus set of the next study and led to continuous refinement of the design of appliance labels. In addition, because the research goals were challenging, it was possible to address them in an organized and systematic fashion. For example, in the early focus groups, participants were shown examples of quite different labels from different countries and were asked to make general statements about these labels. As a result of this broad stimulus array and the general preferences expressed by participants, it was possible to identify favored color schemes (yellow and black was preferred to blue and green), so that the labels employed in subsequent studies had a yellow and black color scheme. Subsequent studies could then focus on other aspects of the label design. In the early studies there was a strong expression of preferences regarding the graphical representation of energy consumption. In all cases, the intent was to demonstrate the relative energy consumption of a particular model relative to similar appliances. Some examples included a thermometer-type gauge, others were more like a car speedometer. Participants showed a clear preference for scales where increasing energy consumption was represented by a left-to-right movement on the scale so this preference was incorporated into subsequent designs.

While the project as initially conceived involved multiple studies, the final study (the shopping experience) was not planned in its final form. This study was the result of flexibility in the research design and budget, as well as a desire on the part of the researchers to conduct a more naturalistic study to gauge the true impact of label design on appliance preference. The shopping experiment was developed in response to concerns that categorical rating systems, particularly the stars-based rating, might mislead consumers by implying product quality in addition to energy efficiency. Thus, the main purpose of the study was to test the impact of each label design on consumer purchase decisions and perceptions of product quality and value. In addition, the study represents the only research task where participants did not focus on the energy labels being tested. By simulating a real-world shopping environment, the study allowed investigation of label comprehension and the role of energy efficiency in

appliance purchasing decisions in the retail environment. Earlier exploration of these issues was limited by the research setting in which the research was conducted. Since participants were not primed in any way regarding energy efficiency, the findings provided a good insight into real world behavior as well as support for the final design recommendations. It should also be noted that while the consumer studies lent themselves to a cumulative design approach, the findings from the study of supply side actors also fed into the final report and recommendation, and are being used to develop the final FTC petition.

Another important feature of the research design was the use of different research methodologies. The early studies of consumers employed a focus group methodology, providing qualitative intensive data for understanding the reasons behind preferences. The later studies then employed a quantitative approach to compare comprehension and understanding of the labels developed from earlier rounds of the projects. In general, an important feature of this multi-method research is the ability for researchers to triangulate their findings by employing different methodologies with independent samples to determine the extent of convergence. The variety of methods that were employed was a major strength of this research. It was clear that different tasks required different research designs. The task was not simply to identify the most comprehensible from an existing set of labels, but to identify the preferred features to develop possibly new variants. Hence the focus groups' and consumer interview findings about label preferences were elaborate and detailed, and helped answer the 'Why' questions about label comprehension. Subsequently, these findings were tested empirically using large-scale survey studies and quantitative analysis.

One feature of the design that was critical involved the rotation of the order of stimulus presentation. Rotation of order is known to be important in determining the efficacy of a range of stimuli, in situations where learning or presentation effects can influence the comprehension and preference for subsequent stimuli. Researchers were careful to vary the order of presentation of various labels in all the studies so that inferences about efficiency could be made. In all the studies except for the final study, the existing Energy Guide label was included as part of the stimulus presentation. This allowed researchers to determine the extent of change (compared to a known baseline) in comprehension and preference that could be attributed to a particular label design.

In such an important project as this, the issue of researcher independence needs to be addressed. Whilst ACEEE was clearly the researcher in this instance, and its agenda on energy matters is well-known, it was considered important to maintain some distance between the participants and ACEEE. To that end, ACEEE employed an independent market research organization, Shugoll Research, to conduct a significant number of the studies.

Conclusions

This research reveals the limited impact that the current EnergyGuide label has on the product choices made by U.S. consumers. Although consumers are familiar with the yellow energy label, most consumers were unable to identify the current label from a group of different label designs, despite the fact that most had recently purchased an appliance or were currently shopping for one. Research findings also provide strong evidence that the EnergyGuide can be redesigned to improve consumer comprehension, encourage wider use of the label, and motivate consumers to consider energy use when purchasing a labeled appliance. In general, an improved label would retain the yellow label format and current EnergyGuide logo; slightly reduce the level of explanatory text; and reposition the ENERGY STAR to the bottom right-hand corner of the label.

More specifically, an improved label would incorporate a categorical rating system. Most consumers preferred categorical ratings over continuous scale graphical designs. Among the categorical

ratings tested, stars emerged as the most preferred and most promising design. This design incorporates the well-recognized, intuitive stars rating and is complementary with the ENERGY STAR. Consumers found the optimized stars label shown in Figure 2 easiest to understand and most motivating. The stars label builds on the familiar yellow EnergyGuide format and incorporates the stars-based rating, preferred ENERGY STAR placement, and an optimized level of explanatory text.

Finally, it should be noted that the significant outcomes of this research were established through the program of a multi-study, multi-method research that allowed significant refining of the label design and the triangulation of findings with independent samples. This type of research is highly recommended for complex, highly visible projects such as this.

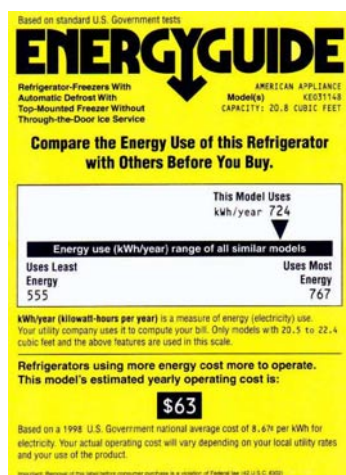


Figure 1: Current Energy Guide

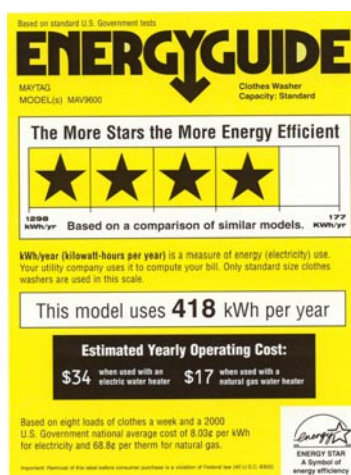


Figure 2: Stars Label

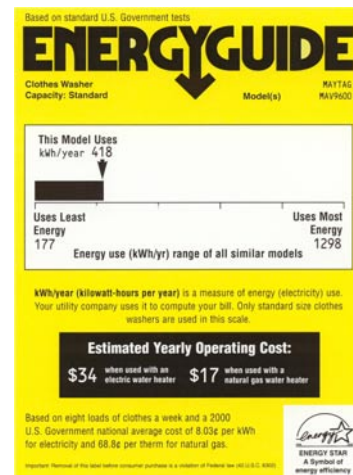


Figure 3: Bar Graph Label

References

- du Pont, P. 1998. "Energy Policy and Consumer Reality: The Role of Energy in the Purchase of Household Appliances in the U.S. and Thailand." Ph.D. Dissertation. Newark, Del.: University of Delaware.
- Thorne, J. and C. Egan, C. (2002). An Evaluation of the Federal Trade Commission's EnergyGuide Appliance Label: Final Report and Recommendations. Washington, DC.: American Council for an Energy-Efficient Economy.