Graphical Displays and Comparative Energy Information: What Do People Understand and Prefer?

Christine EGAN
American Council for an Energy-Efficient Economy

1 - INTRODUCTION

Information policy refers to mechanisms targeting behavior such as labeling, advertising and media campaigns, community outreach, public education programs, direct feedback, etc. Such policies seek to achieve their goals through the provision of knowledge that would not otherwise be easily available or is not already provided by existing information sources (schools, newspapers, etc.). Information programs based upon labeling are in use in a variety of policy areas, e.g. appliance energy consumption and nutrition. However, in the United States (US) few comprehensive studies of how to best design such programs have been completed (duPont 1998). Such comprehensive evaluation and research is essential if effective information policies are to be designed and implemented.

This paper examines how energy consumption data is best communicated to residential consumers, and whether or not comparative energy information motivates consumers to consider conservation and efficiency actions. This paper reports on a systematic test of a range of graph-based displays with actual users to determine which is most easily comprehended and most likely to motivate action. The graphs are a form of feedback[1] and are intended to provide information about their recipients' residential energy use at regular intervals. The graphs are also comparative, anonymously comparing recipients' energy use with that of others in relevant groupings. The key advantage of comparative information is that it allows for an analysis of one's energy use outside the limitations of the individual home, to identify anomalies in one's normal usage patterns.

The research questions addressed are:

(1) Do consumers value comparative information about their energy use?
(2) How can comparative information on energy use relative to others be represented graphically in a way that assures ease of understanding as well as accuracy?
(3) Does self-reported data indicate such information programs stimulate conservation and efficiency in their target populations?

These questions are examined through the context of the author's work on the U.S. Environmental Protection Agency's (EPA) Energy Star[2] Billing[3] program. Energy Star Billing (ESB) is a voluntary program that seeks to stimulate customer-initiated efficiency improvements in households. ESB promotes utility bill-based energy use feedback systems. Within the limits of traditional meter reading and billing systems, the program encourages the development and adoption of bills that compare each residential customer's energy consumption with others in the neighborhood, the utility, or similar comparison groups. Traditionally, US utilities have not provided information that would allow customers to relate their energy use to that of other customers. One of the primary objectives of the ESB program is to make home energy ratings as easy to use as the US automobile miles-per-gallon (MPG) rating, which have proven easy for customers to use and understand (Egan et al. 1996; Hill and Larsen 1990; Pirkey et al. 1982). Utilities who sign an agreement to participate in the ESB program are called "partners." Partner utilities agree to provide its residential customers with graphical displays comparing individuals' electric or gas consumption with that of others in a relevant group, such as homes in one's neighborhood or households of a similar square footage on the monthly bill, on a continuing quarterly or annual basis. This paper summarizes the results of three discreet research tasks undertaken as a part of the work on ESB — semi-structured interviews, a survey, and an implementation case study.

2 - INITIAL DEVELOPMENT

The first task undertaken was to develop preliminary graphical displays for the ESB program and test these concepts through a series of semi-structured interviews. The graphical displays were created with several criteria in mind. First, a good display would be readable and understandable by customers. It would have to not only provide accurate information, but lead to valid customer inferences. A good graphical display would put customers in a position to make more informed and confident energy efficiency improvements in their home, resulting in residential energy savings. Additionally, the display would be indifferent to technology, not reliant on computer models with building component/equipment data requirements. This criterion, along with the "understandable to consumers" requirement, led to use of actual consumption data from the customers' own homes, rather than modeled predictions or per square foot estimates. A final criterion was that the set of displays be flexible, allowing any willing utility to incorporate some variation of the suggested options (Egan et al. 1996).[4] Two of the displays created based on these criteria are shown in Figures 1 and 2.

Results of the initial eight interviews showed that the interviewees were receptive to the idea of their utility providing comparative data on home energy use. Two of the eight interviewees indicated that they kept similar records on their own currently. One interviewee mentioned liking the graphs because, as she said, "although you were often told to save energy or that saving energy was good, nobody ever told you how you or your home was actually doing." Another commented that data comparing one's consumption to one's neighbors' would be more compelling than a simple month-to-month comparison of her household use, noting that she would call the utility "to find out what was wrong" if her consumption regularly compared unfavorably to her neighbors'.

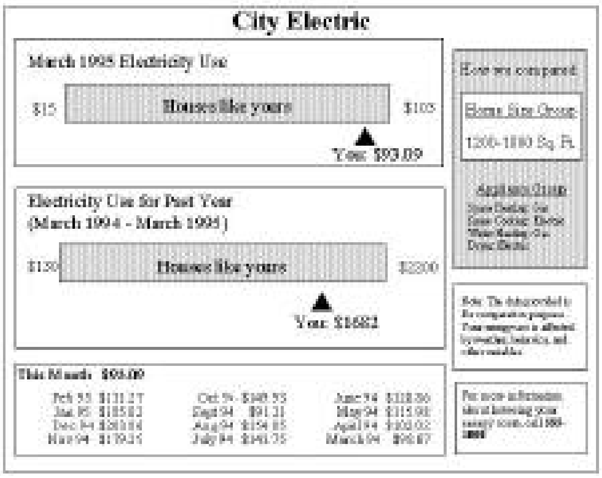


Figure 1 Sample Energy Star Graphic Display

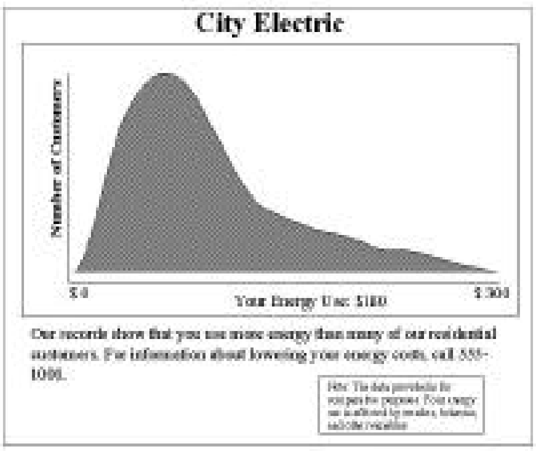


Figure 2 Sample Energy Star Graphic Display

Despite the interviewees' generally positive response, the interviews revealed more misunderstandings of the displays than anticipated. An important and unexpected result that emerged from the interviews was that the bar graph, a central basis for most of the display options, was not automatically and immediately understood by all, or even a majority, of the interviewees. In fact, in four of the eight interviews, significant time and/or explanation was required for the bar graph to be correctly interpreted. In all four of these cases, the problem seemed to be with recognizing that the length of bar represented a range of utility bills for different customers. In one case, the interviewee assumed that the low end was a daily cost and the high end an average monthly cost. The result was that she rated her consumption (represented by the triangle labeled "You") as below the monthly average. Given her interpretation, there would have been no point along the bar graph at which this interviewee would have rated herself as above average. A complete explanation of the bar was necessary to move the interview forward. In another case, the interviewee had trouble understanding that the bar represented a range of costs because the low end was, to her, too low to be realistic. Again, substantial explanation was required to move the interview forward. These results pointed to a need to better refine the test graphs and to continue to observe the rate of understanding not only of the bell curve, but also of the bar graph.

As an initial step in refining the bar graphs, the ESB graphs were compared with those required by the U.S. Federal Trade Commission (FTC) for appliance labeling. The FTC "Energy Guide" appliance labels clearly identify the low and high ends by labeling them "uses least energy" and "uses most energy." It was hypothesized that a similar label, perhaps saying "lowest consumer" and "highest consumer" might enhance interviewees' comprehension of the ESB test graphics. These initial eight interviews (specifically the difficulty 50 percent of our interviewees had in recognizing the bar as a range) point to the importance of such endpoint labels. Further testing was important, however, as it was also possible that the endpoints could further confuse residential bill payers. In the survey discussed next, descriptive endpoint labels were added to the graphs.

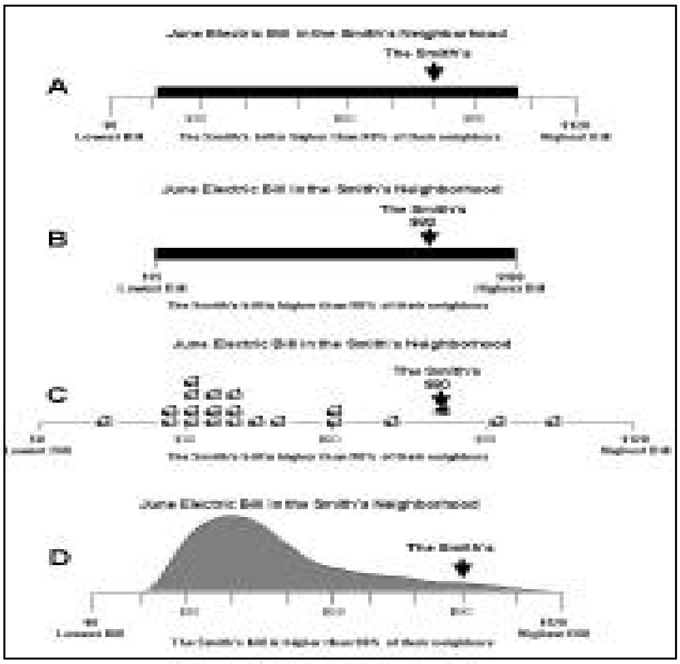
The bell curve (Figure 2) resulted in the strongest reactions from informants. Although in six out of the eight interviews the bell curve was interpreted correctly, most did not react favorably to this display. Displays containing the bar graph were preferred by all interviewees (although this was not true on the larger sample survey reported on next). Six of the eight interviewees commented negatively on the bell curve. When presented with the bell curve, one interviewee who correctly interpreted it exclaimed, "I don't like it, and if I got this in my electric bill I would throw it away." The inclination to throw away the bell curve was mentioned by two other interviewees as well. Another interviewee, who only interpreted the graph correctly after some examination, commented, "I hate these charts, they mean absolutely nothing to me." In addition, two interviewees who understood the bell curve graph suggested that they believed the average person would not. This suggestion by our respondents was not taken as factual; however, such comments (along with the preliminary testing results) pointed to possibly serious problems with the bell curve. It appeared that some people did not understand it, most did not like it, and a couple judged it was too difficult for others. These negative results were expected, and taken along with the surprisingly low comprehension of the bar graphs, were constructive in suggesting the potential value of graphs that conveyed distribution information via alternatives.

In summary, neither of the graphic displays were perfect and further display improvement, along with additional testing, was judged essential. For example, basic improvements were needed in the layout of the displays. Also, with a sample of only eight, a six-to-two or even an eight-to-zero split could not be measured statistically. Thus the graphs were simplified and tested on a larger sample through the survey discussed next.

3 - SECONDARY TESTING

The survey analyzed asked 600 utility bill payers in Delaware to interpret four draft graphical displays and decide their display preferences, comparison group preferences, likelihood of acting on a given display, and willingness to pay for such a service. The four graphics, shown in Figure 3, were developed based on literature review and the pre-testing just discussed.

The four draft displays are of two types _ bar graphs and distribution graphs. All four display the same household (The Smith's) as a comparatively high consumption household. Graph locates the Smith's residence among the range of homes (shown by the bar) in their neighborhood. Graph B, the simple bar graph, does the same, but has only endpoint labels and no scale. Graph C, a distribution graph, shows houses in the Smith's neighborhood mapped according to their monthly bill amounts, with the Smith's house darkened. Graph D is a simple bell curve, with an arrow showing where along this curve (representing the distribution of monthly bills in the Smith's neighborhood) the Smith's home falls. The project team hypothesized that respondents would prefer bar graphs over distribution graphs. Graphs C and D were included among the range of options because of the supplemental distribution information they provide. Prior research using actual utility distributions had pointed to substantial tradeoffs between ease of understanding and completeness in choosing between bar graphs and distribution displays (Egan 1995). Because bar graphs contain no indication of the comparison group layout, they can be misleading. Bar graphs depicting irregular distributions can result in highly inaccurate representations of the underlying data, such as customers in the 80th percentile appearing in the middle of the graph closer to what appears to be the 50th percentile.



The survey instrument was divided into two sections, the first with a single graph and the second section with all four graphs including the one presented in the first section. The graph in the first section was rotated, with one fourth of the sample receiving each possible graph initially. All received the identical set of four. In the first portion of the survey, two questions followed the single, rotated graph: one question testing comprehension, the second testing reported behavioral intention in response to the graph. Next, recipients received the display of all four of the graphs (A through D). The remaining questions evaluated respondents' graph preference, the actions respondents might undertake, comparison group preferences, and willingness to pay for comparative information.

The target population for the full survey was US residential energy bill payers. The project team hoped to reach individuals in homes who pay the utility bill every month, rather than a member of a household who only receives energy services. In the interest of expedience and cost, we limited the survey to homes in Delaware as we had no reason to believe that Delaware bill payers would differ significantly from bill payers in the rest of the country in terms of their interest in the program or ability to comprehend ESB graphical displays.[5] The survey had a return rate of 46 percent.[6] Of the total returns, 28 percent[7] were those sent Graph A (the bar with scale) in the survey's first portion, 25 percent Graph B (the simple bar), 22 percent Graph C (the distribution with houses), and 25 percent Graph D (the simple bell curve). A multinomial distribution test using chai square indicates that the difference between these groups is not statistically significant.

The mean age of survey respondents was 51 and the median 50. Only 0.4 percent of the population was under 20 (one person aged 18). The remainder was approximately evenly divided among ages from 30 through 70, with smaller numbers in the 20s, 70s, and 80s. There was a roughly even distribution of males and females. The sample represented the distribution of educational levels in the population as well.

3.1. Comprehension

When presented initially with the single graph, respondents were asked "to make their best guess about the Smith's electric bills compared with most others in their neighborhood." The majority of respondents (67 percent) correctly interpreted the graphs and answered that "the Smiths would have higher bills." Table 1 shows the percent breakdown of respondents' comprehension by the graph they received first. The percent is also calculated by type (bar versus distribution).

Table 1 Comprehension by Graph and Type[8]

Graph Type	Percentage by Graph		
Bar	A, Bar with Scale	63	62
	B, Simple Bar	59	
Distribution	C, Distribution with Houses	79	73
	D, Bell Curve	67	

As Table 2 shows, the distribution with houses (Graph C) appears to be most easily understood by recipients. The difference between the lowest (the simple bar) and the highest (the distribution with houses) ranking graph was about 20 percent. The second ranking graph (the bell curve) trailed the first ranking graph by about 12 percent. However, an analysis of variance using an F-statistic showed that the difference in rates of comprehension among all graphs was too small to be statistically significant.

One practical question was whether the average person more easily understood bar graphs than distributional graphs (such as the bell curve display and the distribution with houses). As discussed, preliminary results from the semi-structured interviews had suggested bar graphs were not the most easily understood, contrary to our expectations and to the presumptions of many policy analysts. To explore this further, we grouped together the rates of comprehension of the two bar graph displays with those of the two distributional displays. As Table 2 shows, only 62 percent of respondents receiving a bar format correctly interpreted the graph, versus the 73 percent receiving a distribution format. An analysis of variance using an F-statistic showed that difference was significant. More specifically, there was a 94 percent probability (observed significance level of 0.058) that this gap was due to actual difference in rates of comprehension (of bars versus distribution graphs) and not chance.

Another surprising result was that the simple bell curve was ranked higher in comprehension (67 percent) than either of the two bar graph displays. This result is unexpected and one few policymakers and researchers would likely have predicted. It would be interesting to conduct further research to qualitatively evaluate the reasons behind these unexpected answers. But the policy-related conclusion must be that a distributional graph of either form enables more accurate comprehension than the bar graphs used today. This indicates potential comprehension problems with other consumer-targeted, bar-based information programs such as the US Energy Guide labeling program.[9]

3.2. Intent

In order to evaluate intended behavioral change in response to the ESB graph, the survey asked, "If you were living in the Smith's house as shown on the above display, would you take any action to reduce your energy costs based on this information?" Seventy-seven percent of respondents answered "yes;" nine percent answered "no;" and nearly 15 percent said they didn't know. The high number of "yes" responses is promising for energy efficiency; however, it must be interpreted carefully. In all instances, respondents were shown a high user case. The sentence below the graph said that the Smith's bill was higher than 90 percent of their neighbors' bills, and the graph suggested that the Smith's use was unusually high. Also, in general, more people say they will take action on a survey of any kind than will in fact do so. Responses to this question do not mean that 71 percent of bill payers receiving an ESB graph would take action. The responses do, however, indicate that the comparative graph has a strong potential to motivate high users.

In the cross tabulation shown in Table 2, we find those receiving Graph C, the distribution with houses, appear to be most motivated. These respondents said they would take action based on the information contained in the graph more often than with any of the other three graphs. Graph A, the bar with scale, ranked second, but trailed the distribution with houses graph by 9 percent. However, an analysis of variance using a F-statistic showed that the differences between the graph groups was not statistically significant.

Table 2 Intentions by Graph[10]

Graph Received First	Percentage Said They Would Take Action
A, Bar with Scale	77
B, Simple Bar	70
C, Distribution with Houses	86
D, Bell Curve	73

3.3. Graph Preference

The next set of questions referred to the set of four graphs in the second part of the survey. In these questions, respondents answered while viewing all four graphs. We asked a series of questions about which graphs respondents found easy or difficult to understand, and a question about which graph respondents would prefer if they could only receive one.

Table 3 Graph Easiest To Understand

Graph Type	Percentage by Form	
Bar	A, Bar with Scale	21
	B, Simple Bar	25
Distribution	C, Distribution with Houses	38
	D, Bell Curve	16

The distribution with houses was rated easiest to understand by the most respondents. The distribution with houses was selected by 38 percent of respondents as easiest to understand. As shown in Table 3, the two bar graphs ranked very closely to one another but trailed significantly behind the distribution with houses graph. According to a chai square statistic, respondents that reported graph preferences are statistically significant at the 0.05 level.

In addition to selecting the one graph they found easiest to understand, respondents were also asked to check all graphs they found easy to understand and all they found difficult to understand (in two separate questions). Table 4 shows the results of these two questions. Once again, Graph C, the distribution with houses, ranked highest. Graph C was most often selected as easy to understand (77 percent) and least often selected as difficult to understand (17 percent). Another interesting point is that Graph D, the bell curve, despite unexpectedly high comprehension rates, was selected by the fewest respondents as easy to understand (36 percent) and by the most respondents as difficult to understand (58 percent). This result is consistent with the qualitative data gathered in the semi-structured interviews where interviewees indicated that even if they understood the bell curve, they also found it difficult or thought other people would find it to be so.

Table 4 Check All Graphs Difficult/Easy to Understand

Graph	Percentage who Selected as Easy	Percentage who Selected as Difficult
A, Bar with Scale	67	24
B, Simple Bar	62	26
C, Distribution with Houses	77	17
D, Bell Curve	36	58

The final question (asked to help gage respondents' graph preference) inquired which graph respondents would prefer to receive. As shown in Table 5, the distribution with houses graph emerged as the most preferred. It was selected by 35 percent of the survey respondents as the graph they would want to receive if they could only have one. The simple bar, Graph B, ranked second, trailing by 9 percent.

Table 5 Graph Respondents Would Want to Receive

Graph	Percentage Selected
A, Bar with Scale	22
B, Simple Bar	26
C, Distribution with Houses	35
D, Bell Curve	18

In summary, the distribution with houses emerged as the preferred graph based on the answers to all four preference questions. By contrast, the simple bell curve is considered more difficult to understand by recipients than the other three, and is the least preferred when recipients select one graph over others for inclusion in their own bills. This is true despite the fact that the bell curve ranked second only to the distribution with houses in terms of recipients' comprehension.

Why is there a gap between people's reasonably good comprehension of the bell curve, and people's lower self-rating of ease in understanding it or wanting to receive it? Qualitative data from our semi-structured interviews support that bell curves are perceived as difficult. Some interviewees indicated that while they understood the bell curve, they didn't think that the average person would. The semi-structured interviews also indicated that while people may understand bell curves, they do not like them. Several interviewees expressed this directly, stating that they would throw a bell curve display away because they did not like that way of presenting data. This is consistent with the survey recipients' apparent lack of interest in receiving the simple bell curve in their own bills. Another hypothesis, not mentioned in the interviews, is that informants did not understand how a set of houses were corrected into a continuous curve. In this respect, Graph C, the distribution with houses, would be expected to be rated as easier to understand.

3.4. Types of Action Respondents Might Undertake

The survey also asked what sorts of action respondents might take based on receiving an ESB graph. We asked recipients to pick one action they would be most likely to take, and to put a plus (+) sign next to any supplemental actions they might also undertake. A total of 167 out of 257 recipients asked did indicate additional actions. Fifty-eight of these put a plus sign next to one item, 72 next to two items, 30 next to three items, and seven next to four items. The results are shown in Table 6.

These responses have several implications. First, they indicate a low level of recipients who would fail to take action. Less than 10 percent said they would throw the graph away, do nothing, and/or plan to do something but probably not get to it. A negligible number selected any of these as a second choice. Once again, the gap between what people say, and what they will actually do, means that the number of people who would take action is probably lower than indicated in the table. Nonetheless, there is evidence that comparative energy information encourages energy efficiency efforts by high consumption residential utility bill payers.

Behavioral changes were selected most often, both as a first and possible second choice. The only exception to this rule is in the case of home improvements as a possible second choice. While only 14 percent of respondents indicated that home improvements would be the action they were most likely to take based on receiving an ESB graph, 38 percent indicated the improvements were a second action they might undertake.

Table 6 Action Respondents Might Take

Action	Percentage of Respondents Who Checked Graph as Their First Choice	Percentage of Respondents Who Marked with a Plus Sign as Possible Other Choice
Behavioral Changes	44	32
Call Utility	22	26
Other	14	7
Home Improvements	10	38
Throw Away and Do Nothing	5	<1
Plan To Do Something and Not Get To It	4	<1
Purchase New Appliances	<1	21

Another interesting result of this question is that 21 percent of the respondents indicated that their first action based on receiving this information would be to call the utility. Twenty-six percent of the respondents indicated that they might call the utility as a secondary response to receiving this information. Again, we would expect those survey numbers to be inflated above the actual behavioral response, but they do indicate a propensity of 21 percent of those on the high bill end of an ESB graph to call their utility.

3.5. Preferences in Comparison Group

As shown in Table 7, no majority preference emerged in terms of respondents' preferred comparison group.

Table 7 Comparison Group Preferences by Percent

Comparison Group	Percentage
Similar Home (size and appliances)	47
Geographical (nearby houses)	43
Neighborhood	24
1 side of block	4
both sides of block	15
Zip Code	3.6
Whole Electric Company	2.8
Group I Choose	0.4
Don't Know	4

Clearly similar home and geographical comparisons are the most promising options. A chai square test confirms that respondents' preference for these two comparison groups is statistically significant at the 0.05 level. However, it is not clear which of these two is most preferred.

In the practical application of the ESB program, the comparison group selected by participating utilities will have to consider data availability as well as customer comprehension. Although utilities have easy access to geographical groupings (through the use of street addresses and the route books used for meter reading), home size and appliance data typically require supplemental data acquisition. Furthermore, it is possible that geographical comparisons result in greater communications between neighbors on the topic of home energy use than do similar home comparisons in which the recipient cannot easily talk with other people in his/her comparison group. If so, geographical comparisons may offer greater residential energy savings potential. In short, determination of the best choice in comparison group involves several decision criteria in addition to the questions of bill payer comprehension and use investigated here.

3.6. Value Placed on the Information

The willingness of the survey participants to pay for the ESB graph was surprisingly high. The average amount that the survey participants who answered this question were willing to pay was 78 cents per month (11 percent did not answer). A total of 122 survey participants (47 percent) indicated they would pay nothing. Whereas 108 participants (42 percent) were willing to pay something, on average $1.66 per month. Estimates of the one-time cost of implementing the ESB program are less than $1.00 per customer with very little on-going costs (Lord et al. 1996). These numbers suggest that it would be possible to run ESB as a fee-for-service program. Based upon the experience of implementing ESB to date, it appears that most utilities prefer not to charge for the program and to instead take the value to customers in the form of a more positive company image. However, an area that could benefit from additional exploration is the effect, if any, a nominal fee has on consumers' perceptions of the value and importance of comparative energy information.

4 - AN INITIAL IMPLEMENTATION CASE STUDY

Traer Municipal Utilities (TMU) was the first utility to implement EPA's ESB Program. TMU is a small municipal utility serving approximately 1,200 residential and commercial customers. Of these, 580 residential customers (all those living within the city limits) were selected to participate in the program. My colleague Anita Eide and I conducted a post-implementation, on-site, evaluation of TMU's program. Our evaluation included a bill payer analysis, which evaluated the opinions of TMU's residential customers through semi-structured interviews with participants.

TMU implemented the program using comparable consumer groups based upon floor area and major appliances.[11] An introductory letter was sent one month prior to the mailing of the first Energy Star bills alerting residents to the impending program and its purpose.[12] Unfortunately, nearly all of the design and customer analysis work discussed as a part of the survey was completed after the TMU bill was developed. As a result, and despite a high level of customer enthusiasm for the new information, the TMU ESB graph encountered significant comprehension problems. Thus, the TMU program is a case study of program implementation without prior testing. The result was that none of the interviewees showed 100 percent comprehension of the ESB graphic. Following is a discussion of the reasons for those problems, based on interviewees' responses and analysis of the TMU bill format and graph design. Three categories of comprehension problems were identified: (1) labeling problems, (2) existing bill format problems, and (3) program design problems.

Labeling Problems _ Interviewee comprehension was complicated by the lack of explanatory information in the basic graph design, shown in Figure 4. TMU is a combination utility with all of their services charged via a single monthly bill. The ESB graph, however, dealt only with the energy portion of the bill, a fact that was not clearly indicated on the TMU bill or graph. As a result, several respondents tried to locate themselves on the graph using the total bill amount instead of the electricity portion of the bill. In cases where the total bill amount exceeded that of the highest user in the comparison group, interviewees were understandably confused.

Even in cases where the cost of electricity was used to locate individuals' homes on the graph, interviewees often had other labeling-based difficulties in correctly interpreting the graph. For example, the "V," meant to denote the bill payer's home among the range of homes in the comparison, was not explicitly labeled. Graph comprehension suffered as a result. Several respondents were completely unaware of the "V's" significance. Others assumed it represented a benchmark or average consumption level for a house their size. In both cases, interviewees struggled to determine how their homes related to the graph. Based on the results presented in the survey, I would expect that even a simple change like a label indicating "you" or "your home" would have improved Traer residents' comprehension.

```
            Your Home (1641 - 3600 sq ft)

                     V
          +-----+-----+-----+-----+-----+
        $10                            $128
```

Figure 4 TMU Graph Design

Finally, Traer residents were assigned to comparison groups based upon a combination of square footage and appliance mix data from actual homes similar to their own. Although the label above the TMU graph indicates that square footage is used in the comparison, it makes no mention of appliances, nor does it state that the comparison is based on homes in Traer. Most of the customers we interviewed appeared to be unaware of the inclusion of appliances in the comparisons. Others were unaware that they are being compared to a range of actual houses in Traer of comparable size. Some were unaware that the low number and high number referred to the lowest and highest actual consumption by a bill payer within their comparison group. Based on the interview and survey results, clear labeling would likely have alleviated most of the comprehension problems of Traer residents with their ESB graph.

Existing Bill Format Problems _ Interviewees' difficulties in understanding the graph were complicated by problems in reading the existing TMU bill. During the interviews, many customers referred to their entire bill as the "electricity bill," leaving their awareness of the individual services included open to question.[13] Several respondents were clearly unfamiliar with the different items on the bill. For example, several customers appeared to be unaware that the term "RES" denoted electricity consumption in dollars. Simple explanations of the various items on the monthly usage summary, and their relationship to the ESB graph, dramatically improved comprehension, pointing to the value of a pre-implementation inventory of bill comprehension.

Program Design Problems _ TMU used four house size groups. Within each of the four home size groups, bill payers were divided into sub-groups based on their appliance mix. The house size comparison groups had wide ranges in an effort to ensure that all of the appliance sub-groups contained enough bill payers for valid comparisons. However, several comparison groups had only one customer due to their unique appliance mix. Unfortunately, Traer's billing software did not provide for the possibility of single household comparison groups. ESB graphs were generated for all bill payers, regardless of whether or not a valid comparison could be made. Bill payers in single household groups received a graph with their bill amount at both the low and high endpoints. Clearly, such a graph provides no meaningful data to its recipients and respondents who had received such a graph were unsure what to make of the information.

Three types of customer response were found in reaction to these comprehension problems. Some customers tried to resolve the problem themselves. A second group called the utility for clarification. A third group abandoned the effort to understand the graph. However, even with these comprehension problems (or perhaps because of them), all but two of the 18 interviewees were able to recall that a new graph had appeared on their bill. Most strikingly, when asked what they thought about the graph, none of the interviewees made negative comments or said they did not like the graph. Seven of the interviewees' comments were clearly positive. Furthermore, the majority of interviewees were willing to pay some amount for the ESB graph. One of the interviewees did not answer the willingness to pay question. Of the remaining 17, ten said they were willing to pay some amount per month for the ESB graph. Among those ten, the average willingness to pay was 86 cents per month, the lowest amount recorded was ten cents and the highest two dollars; the mode was one dollar and was mentioned four out of ten times. The overall average willingness to pay for the total sample, excluding the one interviewee who did not answer the question, was 54 cents per month.

5 - CONCLUSION

Consumers generally base their energy decisions on information that is easily available, for example, monthly bill amounts or utility bill-based historic self comparisons. However, such tools have significant limitations in that they allow only for comparison within the finite realm of the individual home. Consider the example of an individual living in a poorly insulated home and unaware that his/her residence consumes relatively more than others of a similar size and appliance mix. The traditional monthly energy bill provides this consumer no feedback on his home's relative performance, even if the bill includes historic use data. In short, self comparisons, although an excellent indicator of anomalies in an individual's regular usage patterns (due for example to severe weather or extended absences from the home) are poor indicators of fundamental problems in a residence's overall energy efficiency. The result is consumer energy decisions that are sub-optimal for the home, and given the environmental impacts of energy use, also sub-optimal for society as a whole.

The preceding argument represents the theoretical case for the usefulness of comparative energy information. Through three separate pieces of fieldwork, this paper explored the applied merits of comparative energy information in improving the home energy decision environment. The research provides quantitative as well as qualitative evidence of the value residential customers found in comparative energy information. Participants commented positively on the graphs and indicated a significant willingness to pay. However, comprehension of graphical displays was more complicated than initially anticipated. Surprisingly, non-bell curve, distributional graphs appear to be the most easily comprehended and preferred method of displaying comparative energy information to consumers. This finding is fortuitously congruent with the fact that given the distribution of residential energy consumption, a distribution graph is more likely to lead to valid inferences than a bar graph (Egan 1995). Further research to explore why consumers prefer non-bell curve, distributional graphs would be valuable. The potential of ESB graphs to promote conservation actions is not clearly proven by this research. In part, this is because the tool of comparative energy information is most appropriately tested over a mid- to long-term period. The project team's work on this project, by virtue of its implementation in the initial phases of ESB's development and implementation, could not include long-term testing. A one or two year post-implementation review of the ESB program would offer a concrete measurement of the long-term environmental benefits of comparative energy information.

This research offers several important lessons for policy makers considering labeling and information programs. First and foremost, this research points to the critical importance of clear, well-tested labeling, without which consumer comprehension and program effectiveness are diminished. A more general lesson about information programs is that they may face unique barriers to a thorough evaluation of their effectiveness in that initial screening may overestimate their success. This is because of an apparent discrepancy in consumers' perception of information-providing tools (e.g., their awareness of such tools' existence and their judgement of such tools' merit) and consumers' response to information providing tools (e.g., their comprehension of such tools and their likelihood of taking action based on such tools). For example, in reviewing consumers opinions about the FTC Energy Guide label, duPont and Lord (1996) found that despite a propensity to misinterpret them, consumers were generally aware of the appliance labels and often examined them. This paper is consistent with this finding. For example, in the analysis of the Traer program, despite people's positive impression of the ESB graph, comprehension was very poor. In the survey, using the re-designed graphics, comprehension was markedly improved; however, respondents reported likelihood of taking action was low in spite of the fairly high value placed on the ESB graph. In summary, consumers may indicate a high level of awareness of an information tool, and may even state that they place a high value on that tool. However, this does not necessarily translate to consumer comprehension or a likelihood of using the information provided. Simply stated _ awareness and enthusiasm are not adequate indicators of an information program's efficacy or success.

6 - REFERENCES

duPont, P. 1998. "Energy Policy and Consumer Reality: The Role of Energy in the Purchase of Household Appliances in the U.S. and Thailand." Dissertation for the University of Delaware, Urban Affairs and Public Policy.

du Pont, P. and D. Lord. 1996. "Reality Check: Comparing Policymaker Perceptions with Consumer Energy Behavior." In *Proceedings of the ACEEE 1996 Summer Study on Energy Efficiency in Buildings:* 8.1-8.11 Washington, DC: American Council for an Energy-Efficient Economy.

Dyer, R.F. and T.J. Maronick. 1988. "An Evaluation of Consumer Awareness and Use of Energy Labels in the Purchase of Major Appliances: A Longitudinal Analysis." *Journal of Public Policy and Management* 7: 83-97.

Egan, C. 1995. "Energy Star Billing: A Brief Consideration of Actual Utility Distributions and Graphical Display." Unpublished Manuscript.

Egan, C., W. Kempton, A. Eide, D. Lord, and C. Payne. 1996. "How Customers Interpret and Use Comparative Graphics of Their Energy Use." In *Proceedings of the ACEEE 1996 Summer Study on Energy Efficiency in Buildings:* 8:39-8:45. Washington, DC: American Council for an Energy-Efficient Economy.

Gardner, G.T. and P.C. Stern. 1996. *Environmental Problems and Human Behavior*. Needham Heights, MA: Allyn and Bacon:

Harrigan, M., W. Kempton, and V. Ramakrishna. 1995. *Empowering Customer Energy Choices: A Review of Interaction and Feedback in Energy Efficiency Programs*. Washington DC: Alliance to Save Energy.

Hill, N.W. and R.P. Larsen. 1990. "Evaluation of the Fuel Economy Information Program: Interim Report." Argonne, IL: Argonne National Laboratory.

Lord, D., W. Kempton, S. Rashkin, A. Wilson, C. Egan, A. Eide, M. Iyer, and C. Payne. "Energy Star Billing: Innovative Billing Options for the Residential Sector." In *Proceedings of ACEEE 1996 Summer Study on Energy Efficiency in Buildings:* 2:137-2:142. Washington, DC: American Council for an Energy-Efficient Economy.

Pirkey, D.B., B. McNutt, J. Hempill, and R. Dulla. 1982. "Consumer Response to Fuel Information: Alternative Sources, Uses, and Formats." For Passenger Car Meeting, June 7-10, Troy, MI. Warrendale, PA: SAE

7 - ENDNOTES

[1] Feedback is a form of information provision and is defined by Harrigan et al. (1995: 21) as "information about an ongoing action, which enables the actor to refine the action." Feedback devices allow individuals to teach themselves how to accomplish a particular goal such as saving energy (Gardener and Stern 1996: 83).

[2] Energy Star is a registered trademark of the U.S. Environmental Protection Agency.

[3] Through a cooperative agreement, the University of Delaware (UD) carried out the development, implementation, and evaluation of this program for EPA. While a graduate student, the author served as the lead research assistant on this project for nearly two years and also wrote her Master's thesis on this topic (Egan 1997). The project team included Dr. Willett Kempton, Anita Eide, Christopher Payne, Deirdre Lord, and Maithili Iyer. At various points in this paper, research is referred to that was done collaboratively between the author and other project staff. For example, primarily Ms. Eide and the author conducted the semi-structured interviews with guidance from Dr. Kempton and Mr. Payne. Primarily the author conducted the survey with guidance from Dr. Kempton. The case study was implemented jointly by Ms. Eide and the author with guidance from Dr. Kempton.

[4] The criteria were developed by the UD project staff and Sam Rashkin, Energy Star Billing Project Manager at EPA.

[5] The best sample would have been direct utility customer lists but our initial requests to the local utility were refused. Instead, we used a sample drawn from license holders as provided by the Delaware Department of Motor Vehicles (DMV). We eliminated license holders under the age of 20 from our sample to decrease the sample's bias toward homes with multiple drivers (e.g., homes with several teenagers who do not pay the utility bill). Thus, our sample was drawn from the population of Delaware license holders of age 20 or greater. The survey cover letter asked recipients who were not the electric or gas bill payers to pass the survey on to the person in their home who paid the bills. The survey was sent to recipients on letterhead from the University of Delaware's Center for Energy and Environment. A $1.00 bill was attached to the letter to encourage recipients to fill out and return the questionnaire quickly.

[6] Forty-seven of the 600 mailed could not be delivered due to incorrect address information. This reduced our sample population from 600 to 553. A total of 257 surveys were returned.

[7] All figures are rounded to the nearest whole percentage unit.

[8] Percentages are of those receiving the graph first, thus they do not total 100 percent.

[9] The US Energy Guide label has received surprisingly little evaluation. The author is in the planning stages of a comprehensive evaluation of the label. The last federally sponsored study was conducted by Dyer and Maronick (1986). This study examined only bar and scale-based formats and found bars to be superior. duPont (1998) found that about one-third of respondents were unable to correctly interpret the comparative scale of energy use and were thus unable to use this as a basis for comparison. For example, some (nearly 10%) thought the scale applied to only one model. Others (just over 5%) thought it applied only to the one brand. However, duPont did not test alternative approaches to graph design with US consumers.

[10] Percentages are of those receiving the graph first, thus they do not total 100 percent.

[11] Square footage data was obtained by TMU's staff from the Town Assessors Office and appliance data from customers directly.

[12] The letter also included an appliance mix survey that customers were asked to return.

[13] Given the interviewees' mixed use of the tem "electricity bill," it is possible that the problem of using the total bill amount (versus only the amount for electricity) would have persisted even with a clear label indicating the graph pertained to electricity only (versus the other services charged via the Traer monthly bill).