U.S. Labeling Program Evaluation and Label Redesign Strategies

Presented at the Lessons Learned in Asia: Regional Symposium on Energy Efficiency Standards and Labeling

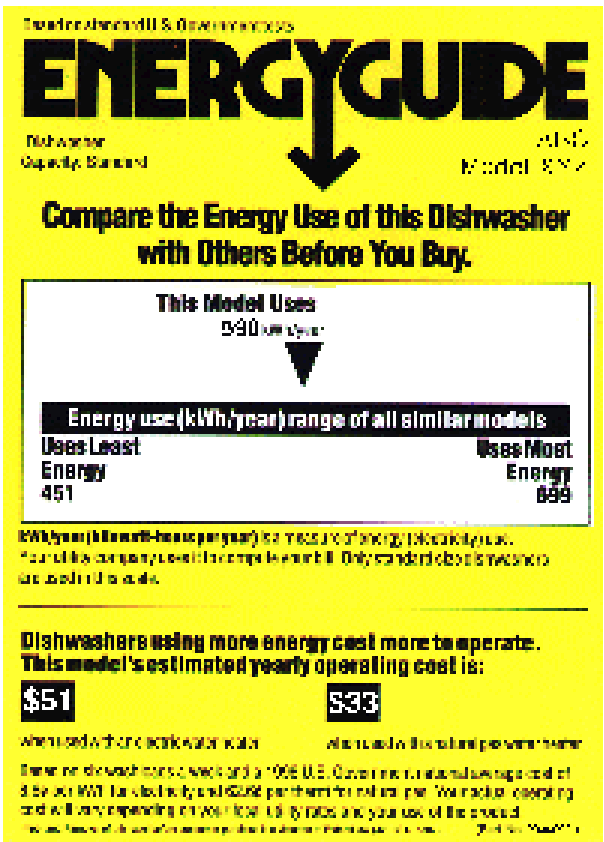
May 29-31, 2001
Bangkok, Thailand

Christine Egan
Executive Director, CLASP

Background

- Energy labels used on appliances in many countries as part of Demand Side Management (DSM) and Market Transformation (MT) programs.
- Two approaches to displaying appliance enrgy performance:
  - Categorical
  - Continuous
- Trend internationally is toward categorical

Current US Label



- U.S. has had a label based upon a continuous scale since 1980.
- Overseen by the U.S. Federal Trade Commission.
- Applied to Products by Manufacturers
- Numbers calculated based upon US DOE Test procedures

CLASP

Project Initiation

- Prior research indicates that U.S. label is not well understood and has limited impact.
- Other research indicates alternative labeling approaches effective where implemented.
- In response, ACEEE initiated an evaluation of the EnergyGuide label in 1999.
- Funded by a mix of utilities, government and NGOS with ACEEE as primary implementer

CLASP

Project Description

- Primary Goal
  - Evaluate effectiveness of current label
    - legislative goal is informational& promotional
- Secondary Objectives
  - Determine best label format for U.S. consumers
    - do bars work better than stars or letters, etc?
  - Prioritize informational elements
    - what is critical and what is clutter to consumers?
  - Uncover opinions of other market actors on efficacy and optimal format/content

CLASP

Research Design & Methods

- Seven graph concepts tested with consumers
  - Current U.S.
  - European-style letters
  - Australian-style stars
  - Thermometer
  - Speedometer
  - Checks
  - Bar with Scale
- Many total label concepts tested making use of these graphs and varying other elements
  - e.g., amount of text, color, etc.

Research Design & Methods (Cont.)

- Initial FGs (6 groups complete 8/99)
  - feedback on labels initial designs
  - emphasized preferences in format/info.elements
  - led to improved designs for further testing
- Semi-structured Interviews (54 interviews complete 9/99)
  - focused on comprehension and interpretation of improved design along w/reasons behind preferences

CLASP

Research Design & Methods (Cont.)

- Second FGs (6 groups complete 2/00)
  - to select optimal designs for quantitative survey
- 3rd FGs (4 groups complete 7/00)
  - to evaluate improved bar-based designs and alternatives to stars (e.g. checks)
  - reactions to inclusion of Energy Star logo

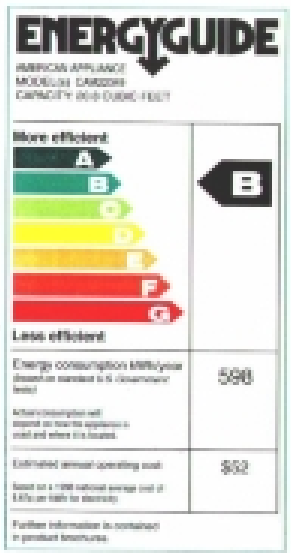
CLASP

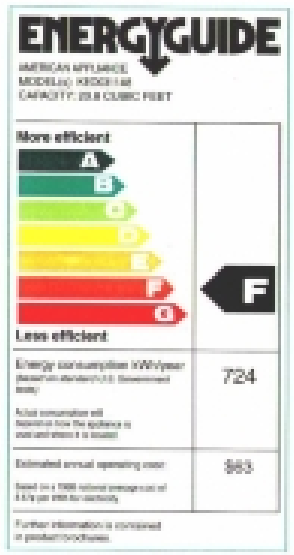
Demand-Side Design

- Survey (n=500, complete 11/00)
  - to quantitatively test comprehension of lead designs along with reported preferences
- Shopping Experiment (05/01)
  - to measure the impact of label improvements on purchasing of efficient models in a simulated shopping setting

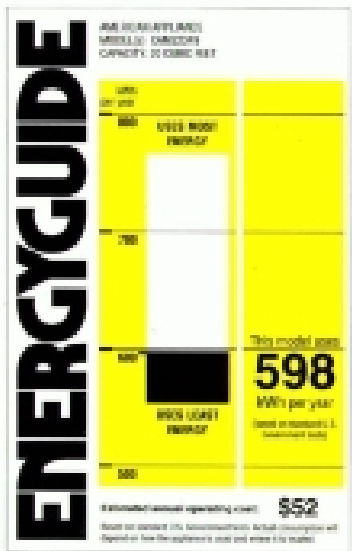
CLASP

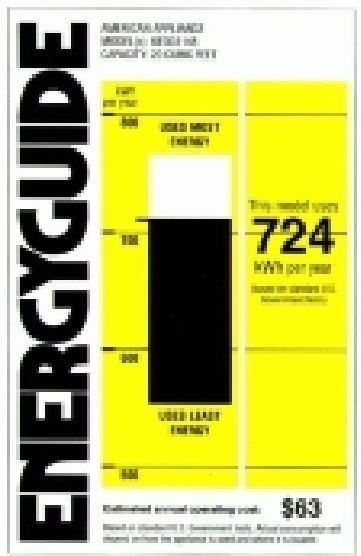
Alternatives Used in 1st FGs

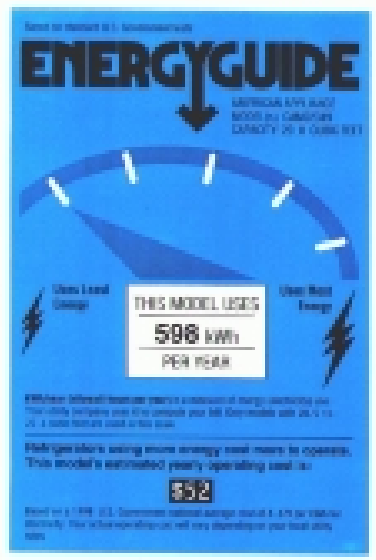


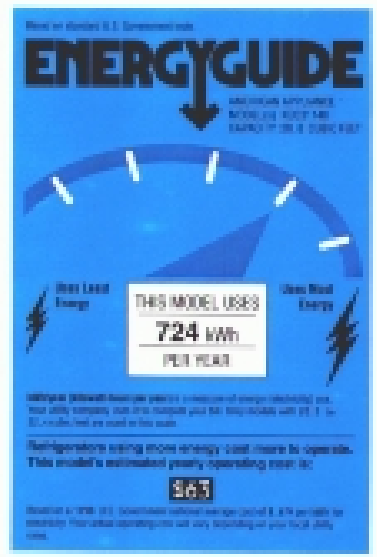


Alternatives Used in 1st FGs









CLASP

Key Findings of 1st FGs

- Low priority on energy efficiency overall
- Current label is familiar but often not used
- Problems with current label design
  - too cluttered
  - poorly organized
  - overly technical
  - graphically unappealing

CLASP

Key Findings of 1st FGs

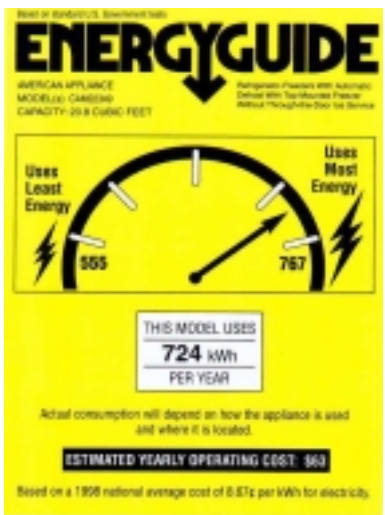
- Ideal EnergyGuide label would:
  - highlight the estimated annual operating and annual kWh so easily seen
  - use the color yellow as a background as associated with energy information
  - use a visually appealing and simple graphic
  - reduce the amount of unnecessary text;
  - clearly state is regulated by the US government
  - use blocked-off spaces and relationally grouped information

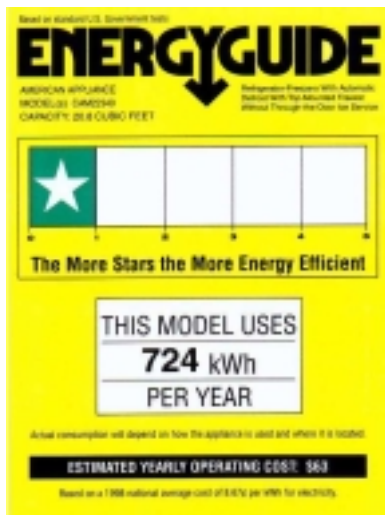
Key Findings of 1st FGs

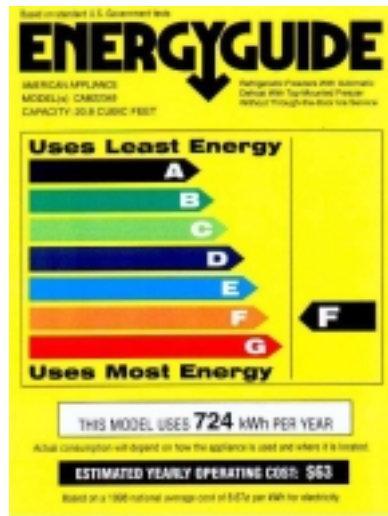
- Letters
  - Longer bars labeled least efficient and shorter bars labeled more efficient is counter-intuitive; exacerbates an existing problem understanding inverse relationship b/t energy use & efficiency
- Stars
  - Negative reaction to lack of operating cost data
- Speedometer/Thermometer
  - Needed graphical design work

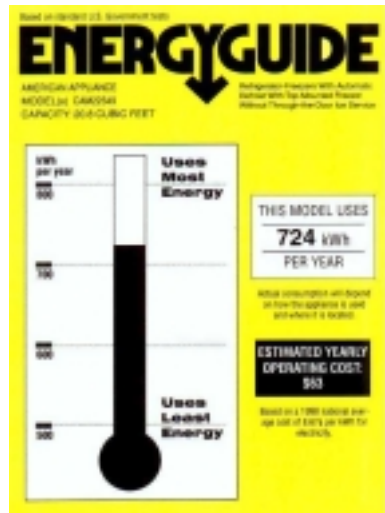
CLASP

Alternatives Used in Interviews









CLASP

Key Findings of Consumer Interviews

- All labels comprehensible to majority but current label appeared most difficult to interpret.
- Overall comprehension problem
  - most people do not immediately grasp the comparative nature of the labels, i.e., this model compared to similar.
- Stars promising b/c of use of intuitive rating scale
- Thermometer & speedometer good but duplicative
- Letters label least refined of all graphic images
  - too many scales of measurement
  - some people want to invert the graph

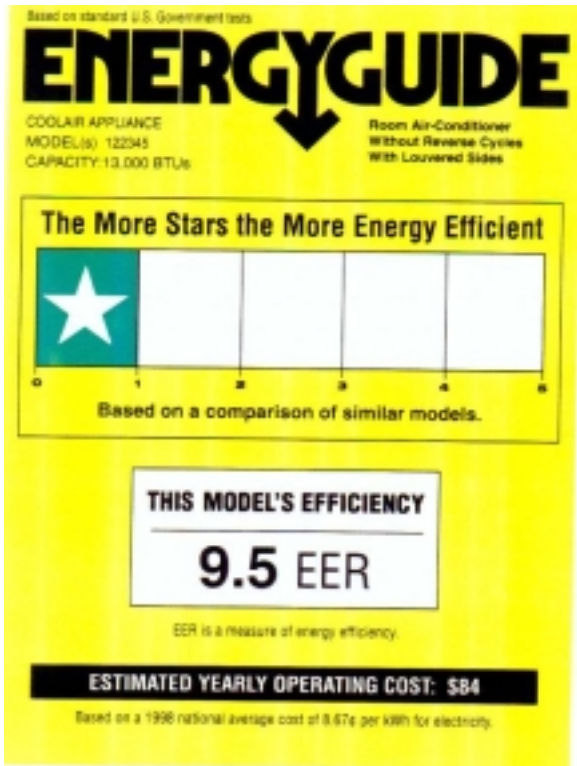
CLASP

Second FGs

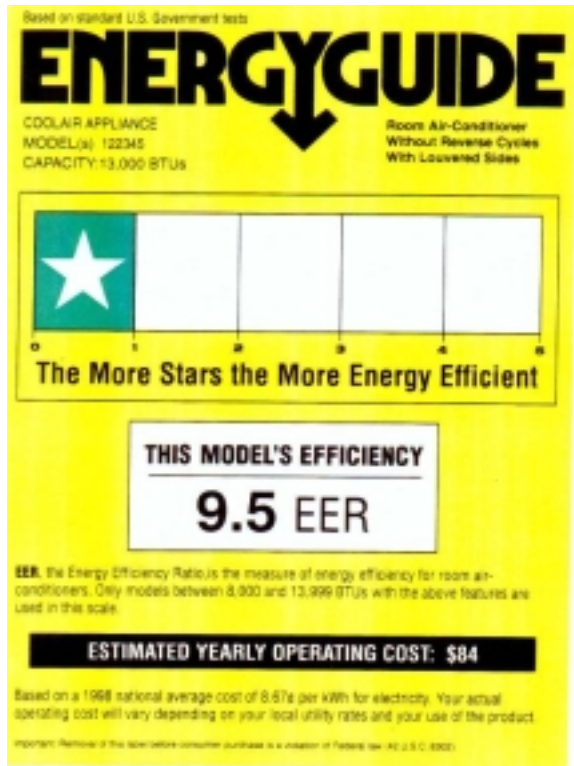
- Speedometer was dropped because it was found to be too similar but not as strong as thermometer
- Multiple versions of remaining designs were developed to incorporate various improvements and suggestions.
  - Versions of stars/letters w/kWh range endpoints
  - Versions of all graphs with low vs. high verbiage levels
  - Versions of letters varying color scheme
  - Versions of thermometer based upon kWh vs.EER

CLASP

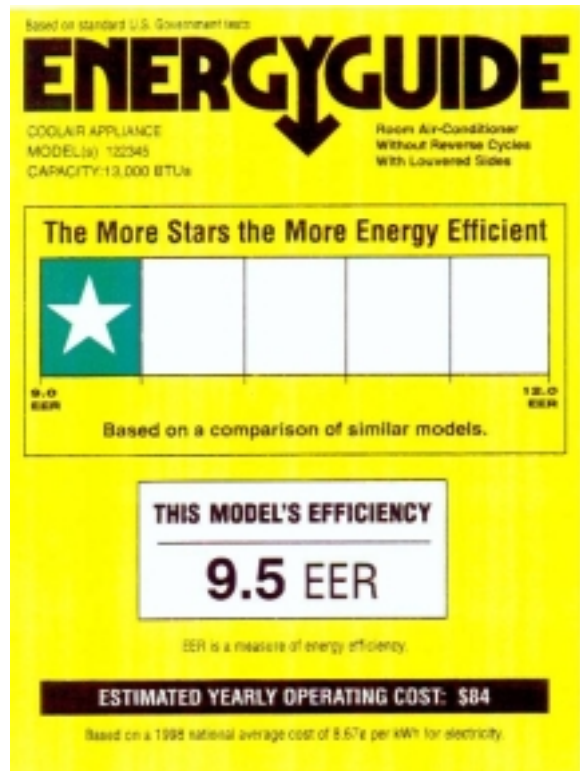
Alternatives Used in 3rd FGs



LOW VERBIAGE



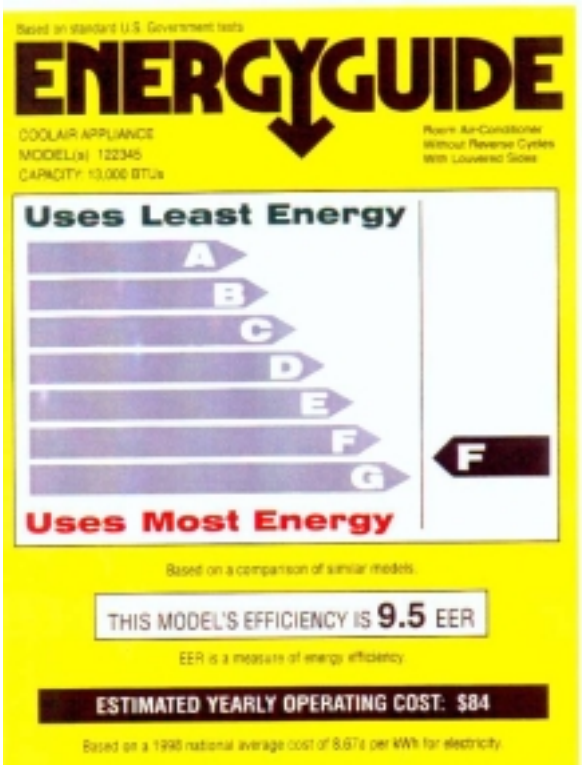
HIGH VERBIAGE



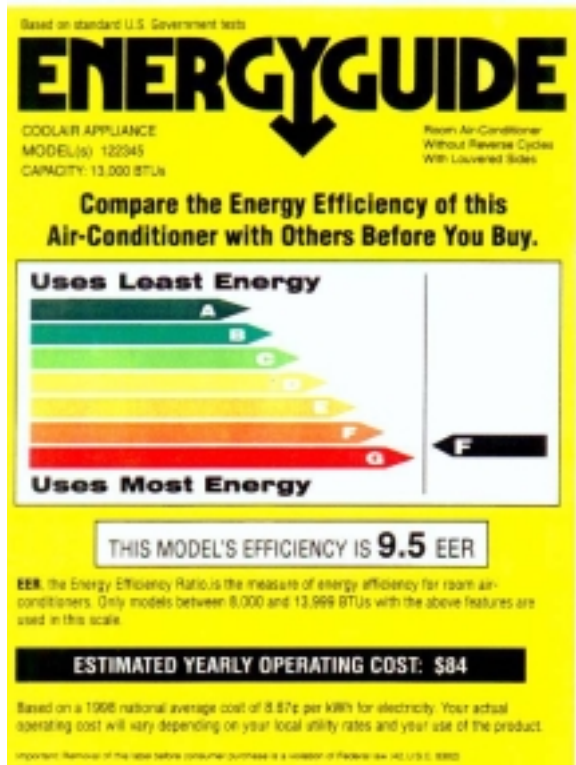
RANGE POINTS

CLASP

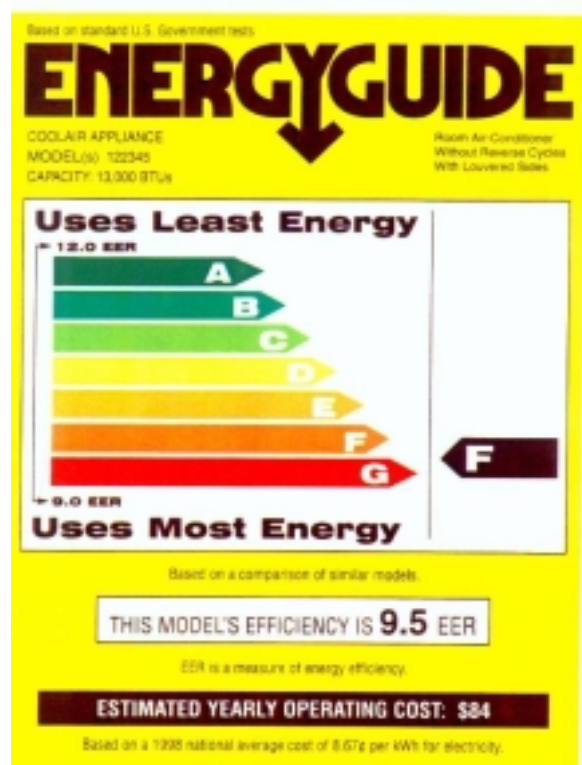
Alternatives Used in 3rd FGs



MINIMAL COLOR



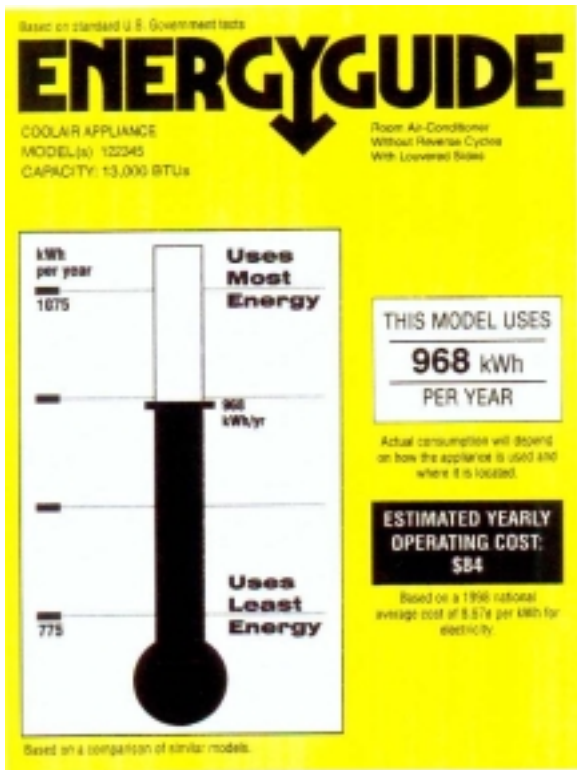
HIGH VERBIAGE



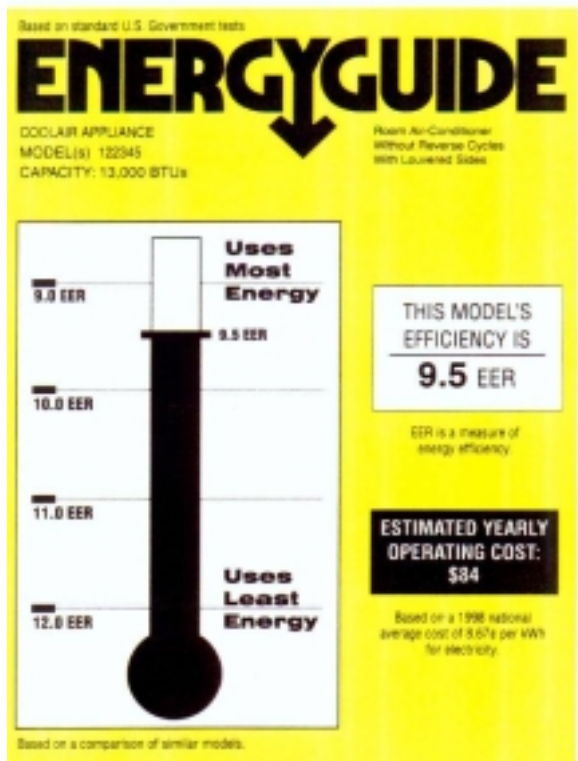
RANGE POINTS

CLASP

Alternatives Used in 3rd FGs



kWh Scale



EER Scale

CLASP

Key Findings of 2nd FGs

- Stars
  - most preferred b/c rating system skips some analytical steps and is eye-catching.
  - Versions w/more information (kWh range and high verbiage) were preferred
- Current
  - second most preferred b/c familiar and has depth of information.
  - But respondents still indicate don't typically read it.

Key Findings of 2nd FGs

- **Letters**
  - **Very-attention grabbing b/c of color scheme**
  - **But serious interpretive problems.**
    - Longer bars meaning less efficient still problematic
    - Some want scale inverted
    - Don't like when color isn't present but spend a lot of time analyzing meaning of color w/few appreciating symbolism of red=stop/bad and green=conservation.
- **Thermometer**
  - **Very negative response**
    - want scale inverted with most efficient product at the top

Key Findings of 3rd FGs

- Alternatives to stars not as attention-grabbing but shouldn't be discounted.
- Stars work well in part b/c some people may relate it to quality
- Energy Star logo (an endorsement label for best-in-class products sponsored by U.S.DOE and U.S.EPA) works well as part of either continuous or categorical labels but placement outside of is crucial

CLASP

Survey

? 10 label executions tested; 5 designs, each presented with and without the Energy Star logo

– Categorical design using letters - J

– Categorical design using checks - K

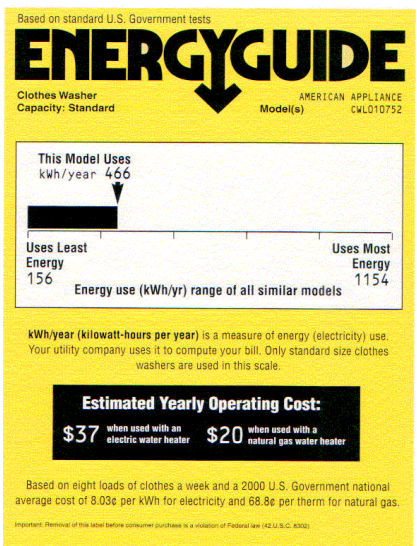
– Categorical design using stars - L

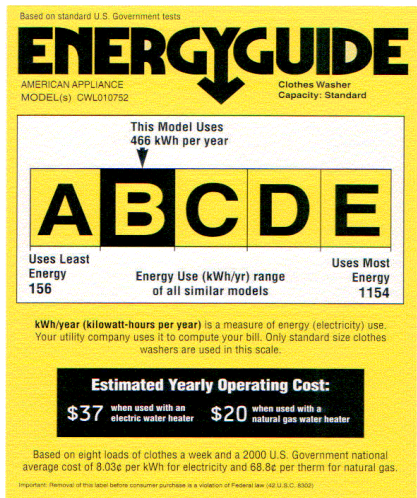
– Current label - M

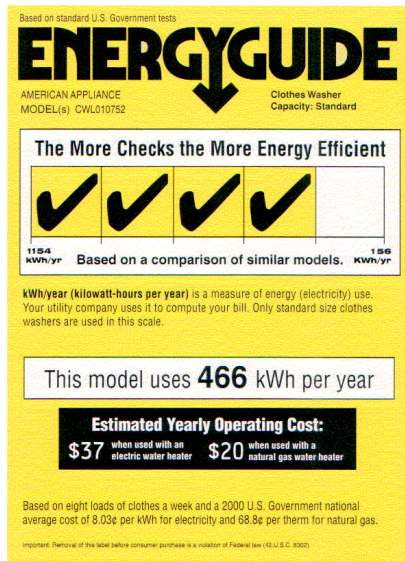
– Continuous label using bar with shaded line - N

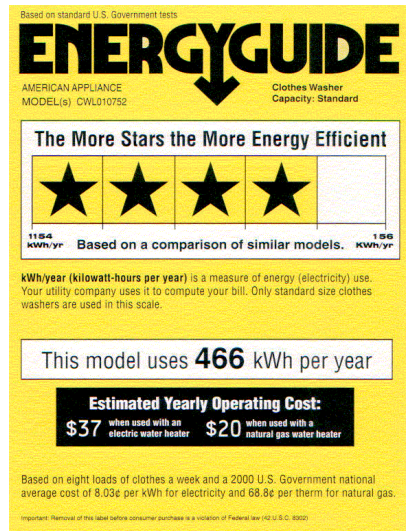
CLASP

Alternative Graphics Used in Survey



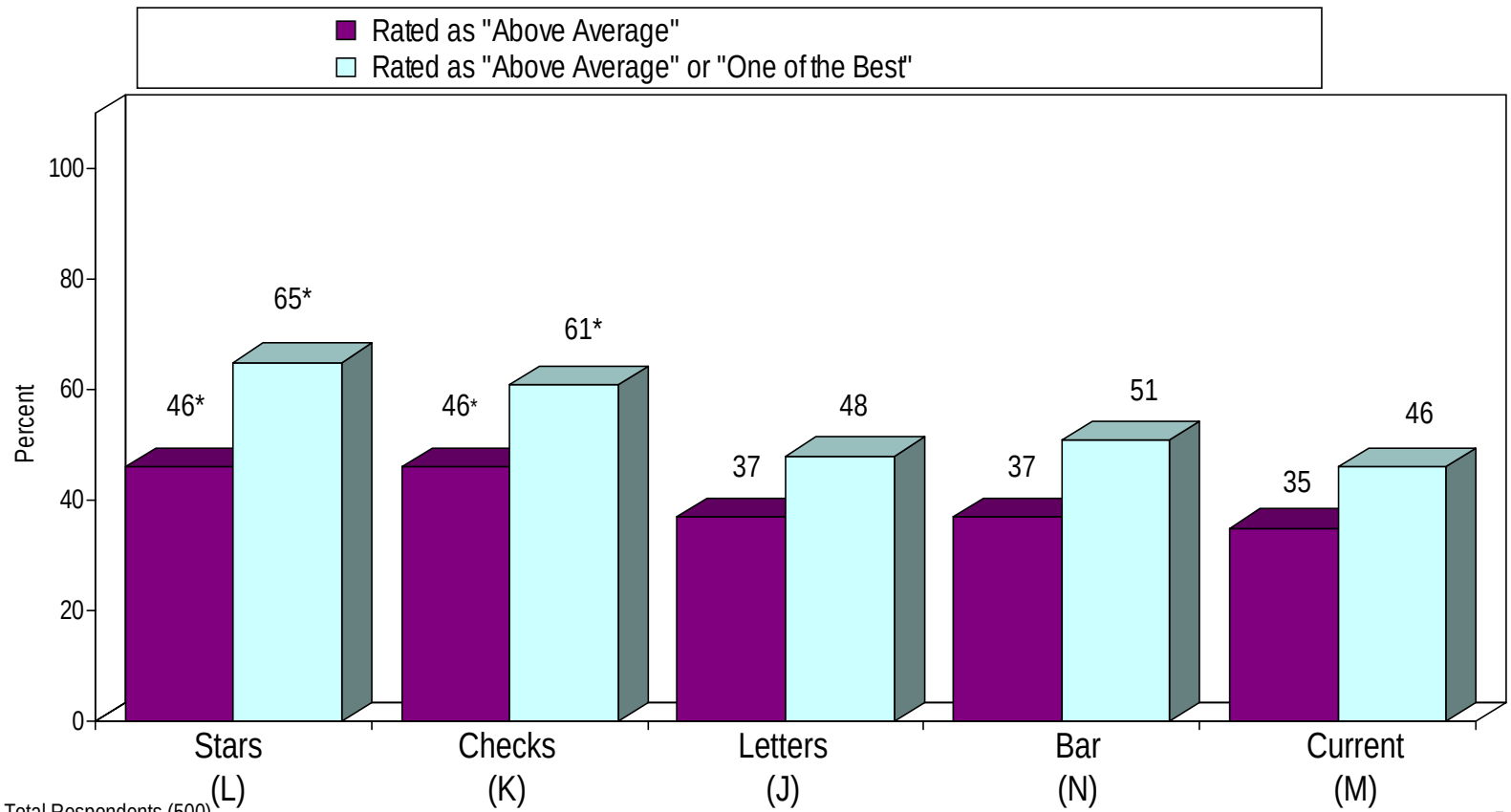






CLASP

Comprehension Results



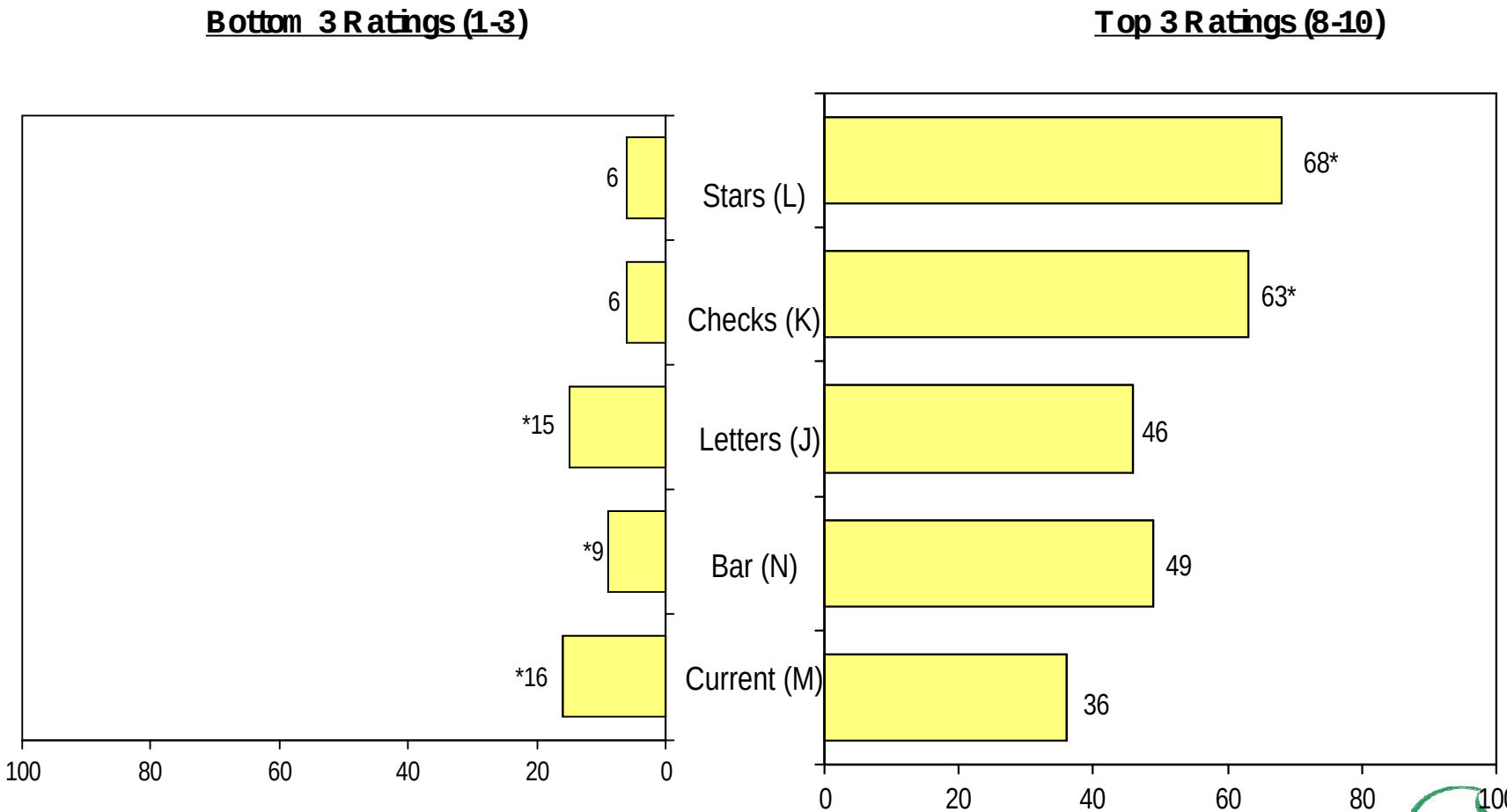
Base: Total Respondents (500)

Note: Percentages shown are possible correct answers to the questions, "Based on the information contained in this label, how would you say the clothes washer that this label describes is in the terms of energy efficiency?" Choices included: one of the best, above average, about average, below average or one of the worst

*Denotes percentages significantly higher than all non-starred percentages at 95 percent confidence level

CLASP

Ease of Understanding Results

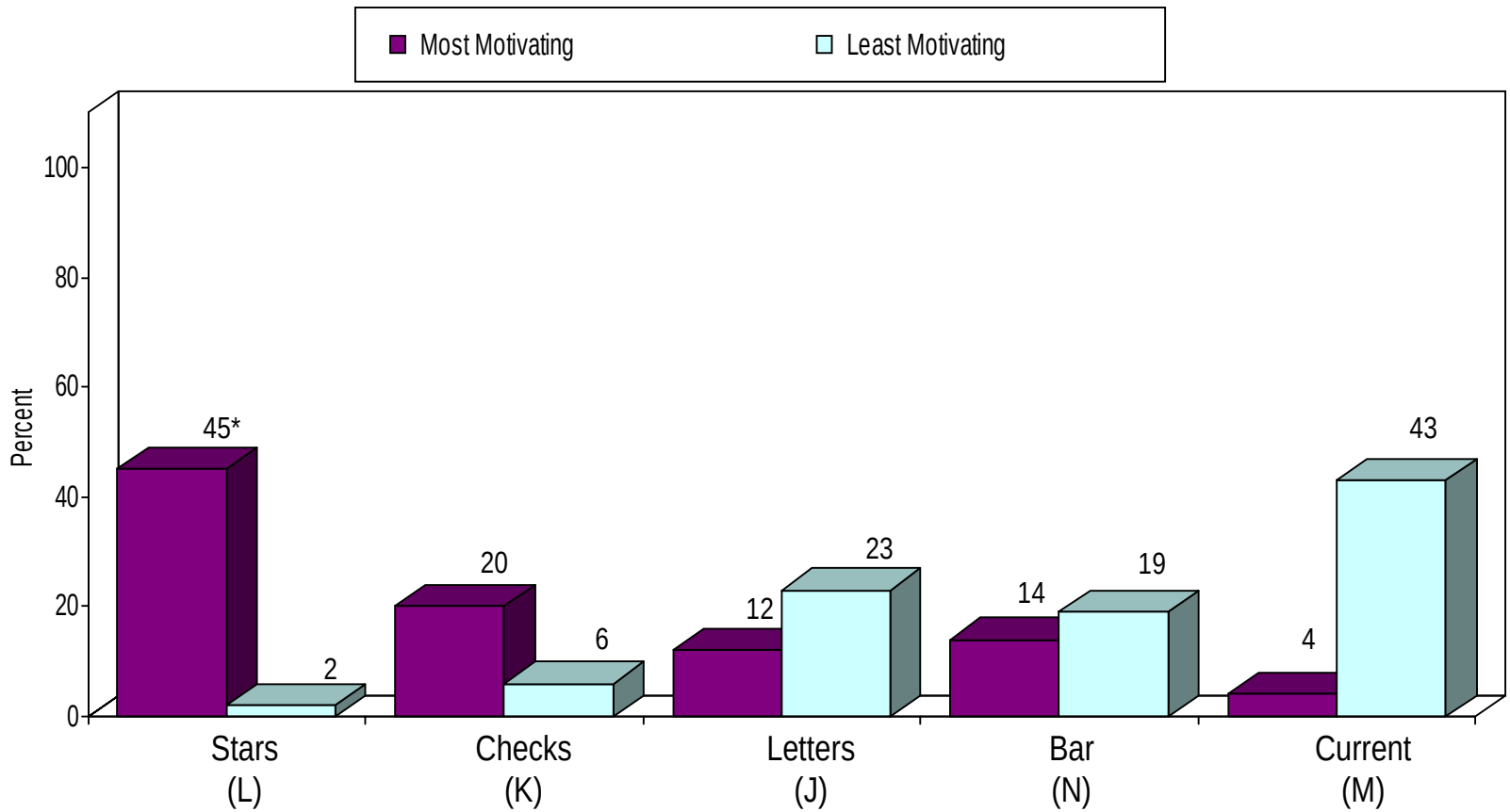


Base: Total respondents (500)

Note: Ratings based on a 10-point scale where 1 represents "Not at all" and 1 represents "Extremely"

* Denotes percentages significantly higher than all non-starred percentages at 95 percent confidence level. ANOVA indicates significant differences in mean ratings at 95 percent confidence level (66.69).

Motivating Ability Results



Base: Total Respondents (500)
Chi-square indicates significant difference across labels at 95 percent significance level (X^2= 253.43/4 d.f.)

CLASP

Summary Survey Results

	Stars (L)	Checks (K)	Letters (J)	Bar (N)	Current (M)
Evaluation of appliance's energy efficiency based on label shown	☙	☙			
Perception of appliance quality based on label shown				☙	☙
Ability to identify most and least energy efficient model from set of 3 labels	☙	☙			
Ease of understanding (Likert scale)	☙	☙			
Right amount of information (Liket scale)	☙	☙			
Forced choice selection of one that best communicates energy efficiency level of appliances	☙				
Believability and credibility (Likert scale)	☙	☙			
Ability to grab attention (Likert scale)	☙	☙			
Forced choice selection of one most likely to read	☙				
Makes you consider energy use in purchase decision (Likert scale)	☙	☙			
Forced choice selection of one that most motivates to consider energy use in appliance purchase	☙				

CLASP

Conclusions

- Results suggest that categorical systems better than continuous at communicating energy efficiency.
  - For U.S. Consumers, Stars are best.
- Primary lesson learned is that iterative, multi-method consumer research is the ONLY way to ensure good label design.

CLASP