# *From Training to Credible Energy Savings: An Evaluation Approach*

***American Evaluation Association Conference***
***Baltimore, MD***
***November 10, 2007***

**Harley Barnes, Ph.D.**
***Energy & Environmental Services***
***Lockheed Martin***
**Presented by**
**John Reed, Ph.D.**
***Innovologie***

LOCKHEED MARTIN

Work Sponsored by

U.S. Department of Energy's

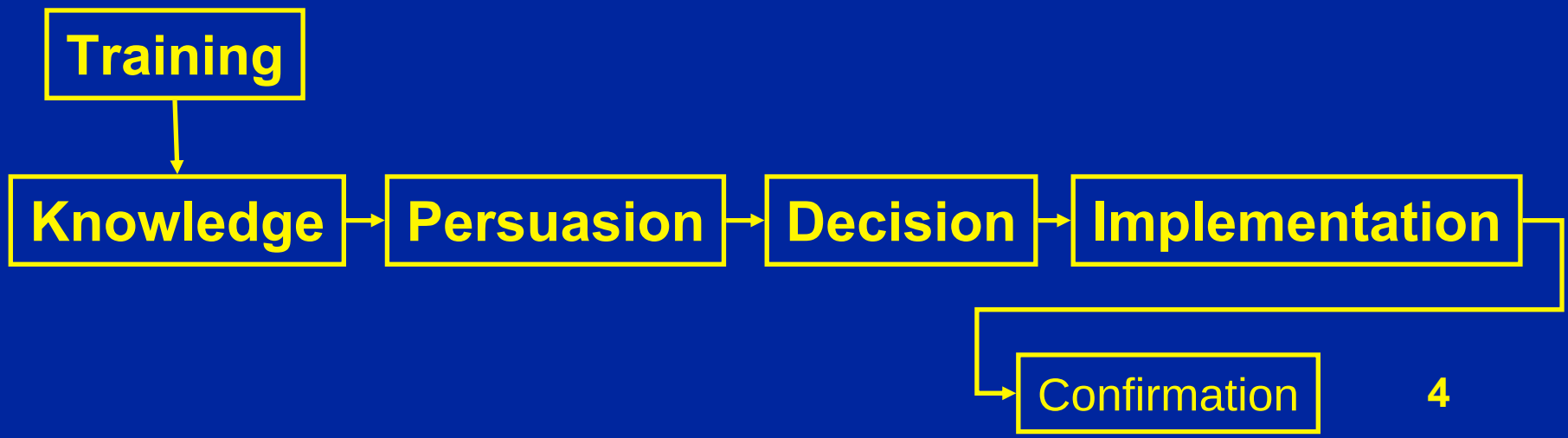
Office of Energy Efficiency and Renewable Energy (EERE)

# General EERE Program Evaluation Goals

- Accountability
  - Was intervention cost effective?
  - Use impact (summative) evaluation
- Program improvement
  - Can we more effectively and efficiently influence decisions?
  - Use process (formative) evaluation
- Manage evaluation costs

# Program Being Evaluated

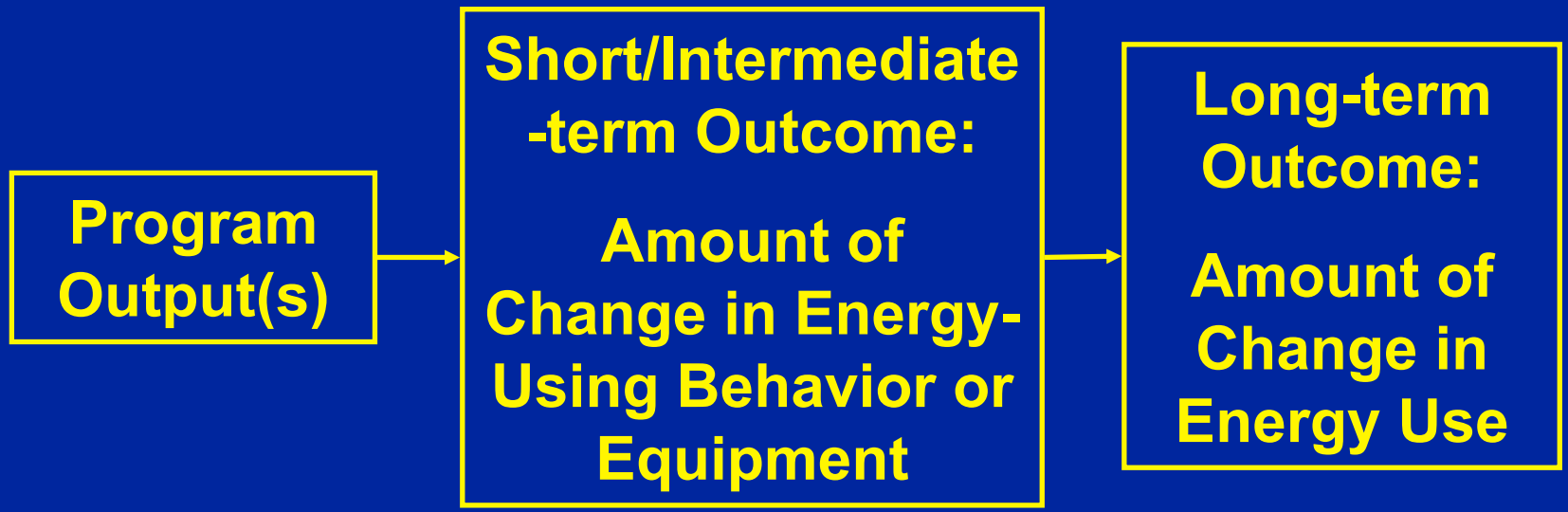
- EERE Industrial Technologies Program, Best Practices End-User Training Subprogram
- Training for industry in actions (measures) that will improve energy efficiency of steam usage
- Has long-term goal of reducing energy usage per unit of production
- High-level program theory:



# Impact Evaluation Objective

- Estimate the energy-savings outcome attributable to training
- If can't measure energy-savings outcome directly, need to measure actions designed to save energy

# Nature of Outcomes for Energy-Efficiency Programs

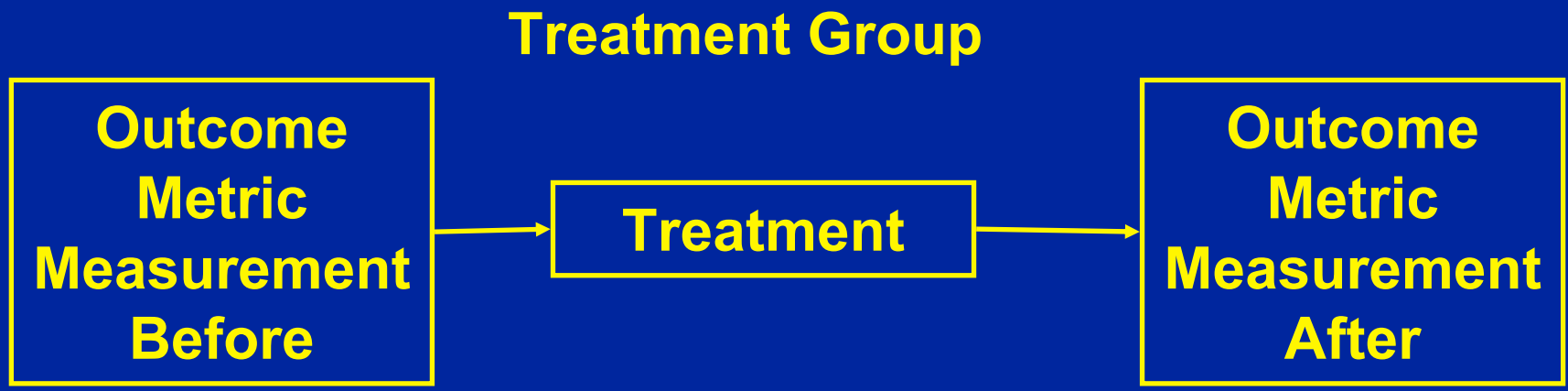



- Logic models are essential

# Energy Program Evaluation Design Challenges

- Relatively small, highly varied population
- Random selection not an option
- Training does not directly lead to long-term goal
  - Multiple additional internal and external influences on outcome
- Availability of credible comparison groups not assured

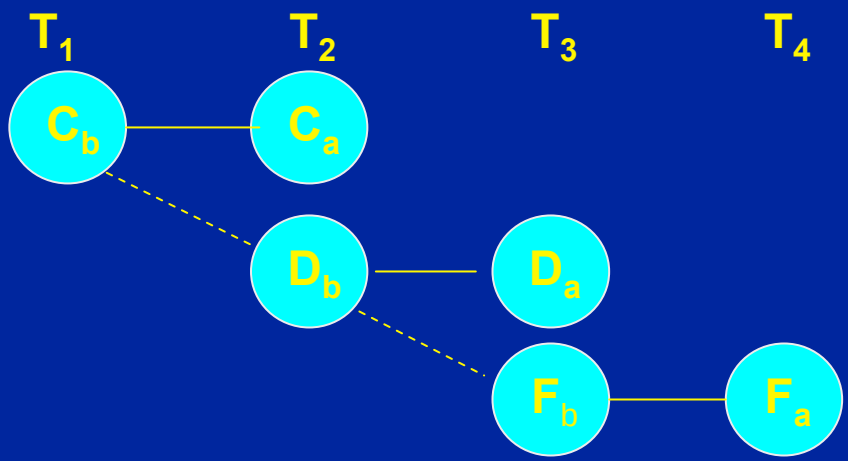
# Classical Evaluation Design



# Level of Analysis for Impact Evaluation

- Classical
  - Measure differential change in group
    - Change in group test scores per class
    - Change in bushels of wheat per acre
  - One outcome metric
  - Group for training consists of all plants represented in training

# Use of Lagged Comparison Group




$C_e$ or $C_i = (C_b - C_a) - (C_b - D_b)_{adj}$

$D_e$ or $D_i = (D_b - D_a) - (D_b - F_b)_{adj}$

T represents time
C, D, and F represent groups of firms
"b" and "a" represent before and after measurements
"e" and "I" represent market effects and impacts
Solid lines represent panels and dotted lines cross section comparisons

# Lagged Comparison Group Design Strengths

- Self-selection controlled
- Can compare what comparison group did during treatment group impact period
  - Assume history affects both the same

# Lagged Comparison Group Design Issues

- Success depends on who volunteers for subsequent training
  - Types of plants
- Judgment whether a matched group exists
  - Pre-established criteria useful
- Relies on comparison group's recall of what did during impact interval
  - Differential outcome comparison will be based on actions, not energy usage

# What to Measure?

- If can't measure energy usage diredtly, have to measure "actions"
  - Short/intermediate-term outcomes
- Training promotes multiple "actions" to save energy in each facility, e.g.,
  - Behavior, e.g. maintenance practices
  - Capital investment in equipment
- What metrics to use for actions?
- Can we use group level approach if we have to measure actions?

# Level of Analysis for Impact Evaluation

- If have to measure short/intermediate outcomes (actions):
  - Can be multiple short/intermediate outcomes in facility
  - Need to measure attribution for these many actions
  - Can we evaluate impact of each?
  - Can't afford to measure each using group level of analysis
- This requires measurement of attribution at the unit (facility) level of analysis rather than group
- Leads to using self-report

# Measurements Using Self Report

- Can ask (and do ask) about overall energy-savings achieved
  - But attribution still relies on actions taken
  - Respondent will think of influence in terms of specific actions
- Survey the trainees about what their plant did after their training

# What Actions Taken

- Ask what energy-efficient improvements made during impact period
  - Open or closed-ended responses
  - Need engineering experts to understand description of what was implemented
  - Could be maintenance or operational changes
  - Could be capital investment
- Ask about quantity of each improvement

# Attribution of Actions

- Try to measure
  - Intent to implement action
  - Timing of implementation
- Use multiple questions
- Use rating scale
  - Simple Yes/No question considered inaccurate measure

# Attribution Options

- Probabilistic attribution result
  - Results in "partial" attribution
  - Outcome is attributed fractionally, e.g., 50% of resulting energy savings is attributable
- Deterministic attribution result
  - Yes/No
  - 100% of outcome or 0% of outcome
- Raises question: Is "impact" a degree-of-influence or deterministic measurement?
  - Usually go with degree of influence

# Degree-of-influence Measurement Philosophy

- Few respondents can say with complete credibility whether they would have done something when they did it solely because of a specific event
- Therefore, use degree-of-influence
- Degree of influence of training on what they did requires
  - Decision theory
  - Question language
  - Question phrasing

# Logic Models Essential to Use Self Report

- Industrial population investment decisions typically consist of several decisions by different roles made in series
  - ► A + B + C = Decision
- Identify business roles and factors affecting decisions on energy use, e.g.,
  - Information flows between roles
  - Role influence
  - Is there a tipping point? If so, where?
- Helps select metrics at different points of influence
  - Helps manage cost of evaluation

# How We Measure Attribution

- Use a scale with corroborating questions
  - Often 0 to 3, single-pole scale
- Use corroborating questions applied to scale to confirm or modify response
  - Requires a “scoring” protocol to assign probability scores.
  - Protocol requires faith in scorer’s judgment

# Wrap up

- Issues and choices described here are universal in the field of energy-program evaluation
- Department of Energy moving to conduct more evaluations of energy programs
- Self-report widely used for accountability and improvement objectives in energy program field
- Can be used with impact calculator to reduce cost of repeated evaluations

Harley Barnes
Lockheed Martin Business Solutions
2277 Research Boulevard
Rockville, MD 20850
(301) 510-6322
harley.h.barnes@lmco.com