ASSESSING CUSTOMER SOLUTIONS AT THE INTERSECTION OF BEHAVIOR AND TECHNOLOGY

Carol Yin, Ph.D., Yinsight
Edwin Hornquist, Southern California Edison

Presented at the Behavior, Energy and Climate Change Conference
Washington, D.C., December 9, 2014

The Story:

- Behavior in CA
- Defining behavior
- NOT defining behavior
- EM&V implications
- What social scientists can do
- What utility program managers can do

Growing Interest in Behavioral Programs

- California investor-owned utility landscape
  - California investor-owned utilities are required to have 5% of residential customers participate in behavior programs.
  - Defines “behavior programs” only as “energy usage disclosure programs”
  - To only be evaluated using “experimental design”.
    - Random selection of test and control groups
    - Energy savings can be claimed on an ex poste basis.

Expanding the Definition of Behavior

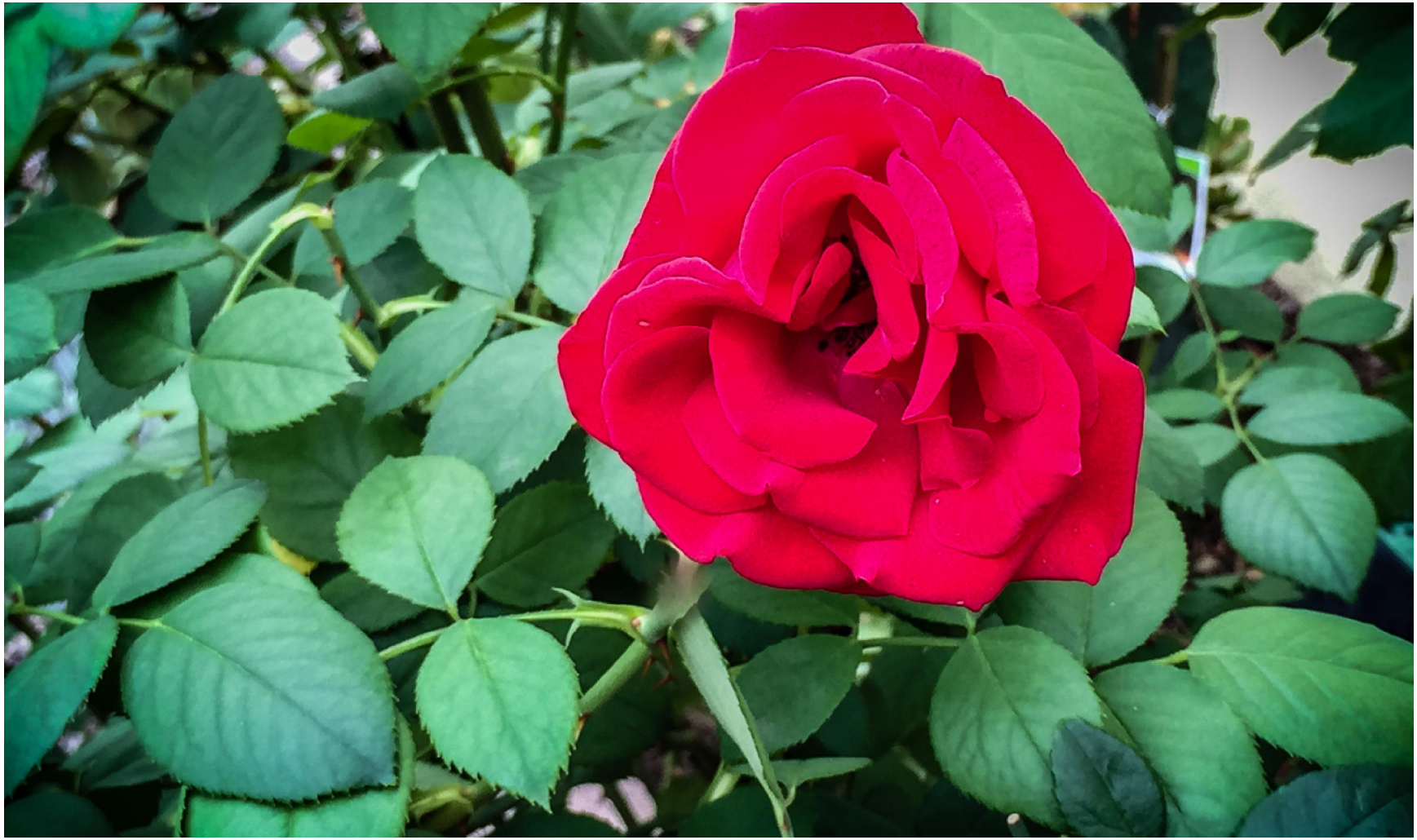
- CPUC and IOUs are collaborating on expanding the definition of “behavior program”

- Must also define and justify EM&V protocols if not using RCT

What is a behavioral program?

- Theoretically-driven definitions
  - Recent whitepapers on behavior
    - Ignelzi et al (2013)
    - Mazur-Stommen & Farley (2013)
- Taxonomy of behavior programs in the field
  - Cognition, Calculus, Social Interactions
  - 15% of “behavior programs” were not behavior-based (Mazur-Stommen & Farley, 2013)
    - “The remainder were technology-based programs, market transformation activities, ENERGY STAR-related programs, and programs with behavioral labels where we could not find a description or be sure they took place”

Leaving Behavior Undefined



Functional Approach to Assessing Ideas for Behavioral Interventions

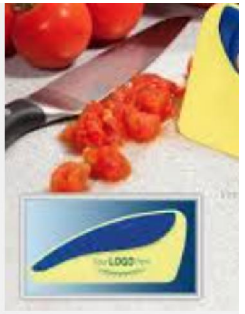
	New behavioral program idea	New technology
Against existing behaviors	Gamification, competition, rewards	Shower timers, dish squeegee
Against existing equipment	BOC, human factors, usability, training	(Yawn)

Can you deem water/energy savings with this?

What's the EUL for this?













Measure against multiple control conditions

- Single variable experiment: Against nothing
- Against biases, placebo effect / Hawthorne effect

Measure multiple test conditions

- Multiple comparison conditions?
  - Compare Program A, Program B, Program C…
- Compare multiple segments!
  - Intervention A in Segment 1, Segment 2, Segment 3
- What does this look like…?

Latin Square

- Counterbalanced Quasi-experimental Design

	Program A	Program B	Program C	Program D
Segment 1	Q1	Q2	Q3	Q4
Segment 2	Q2	Q4	Q1	Q3
Segment 3	Q3	Q1	Q4	Q2
Segment 4	Q4	Q3	Q2	Q1

Full vs Quasi-Experimental Designs

- Full Experimental Design
  - Pro: “gold standard”
  - Con: Expensive, need representative sample
  - Con: Most vendors do not understand research methods
- But, why this gold standard, and not others?
  - Type I vs Type II
  - Replication

That feeling of hopelessness...

- So many sources of “invalidity” to avoid
- Good if these worries lead to design of better experiments

- “It is, however, an unwanted side effect ***if it creates a feeling of hopelessness with regard to achieving experimental control*** and leads to the abandonment of such efforts in favor of even more informal methods of investigation. Furthermore, this formidable list of sources of invalidity might, with even more likelihood, reduce willingness to undertake quasi-experimental designs, designs in which from the outset it can be seen that full experimental control is lacking.” – Campbell and Stanley (1963)

How do the social sciences deal with this?

- “From the standpoint of the final interpretation of an experiment …every experiment is imperfect.” – C&S (1963)

- Can’t prove a hypothesis, can only try to rule out alternative explanations of your findings.
- Quasi-experiments can certainly help in ruling out alternative explanations.

Benefits of using Quasi-Experimental Designs

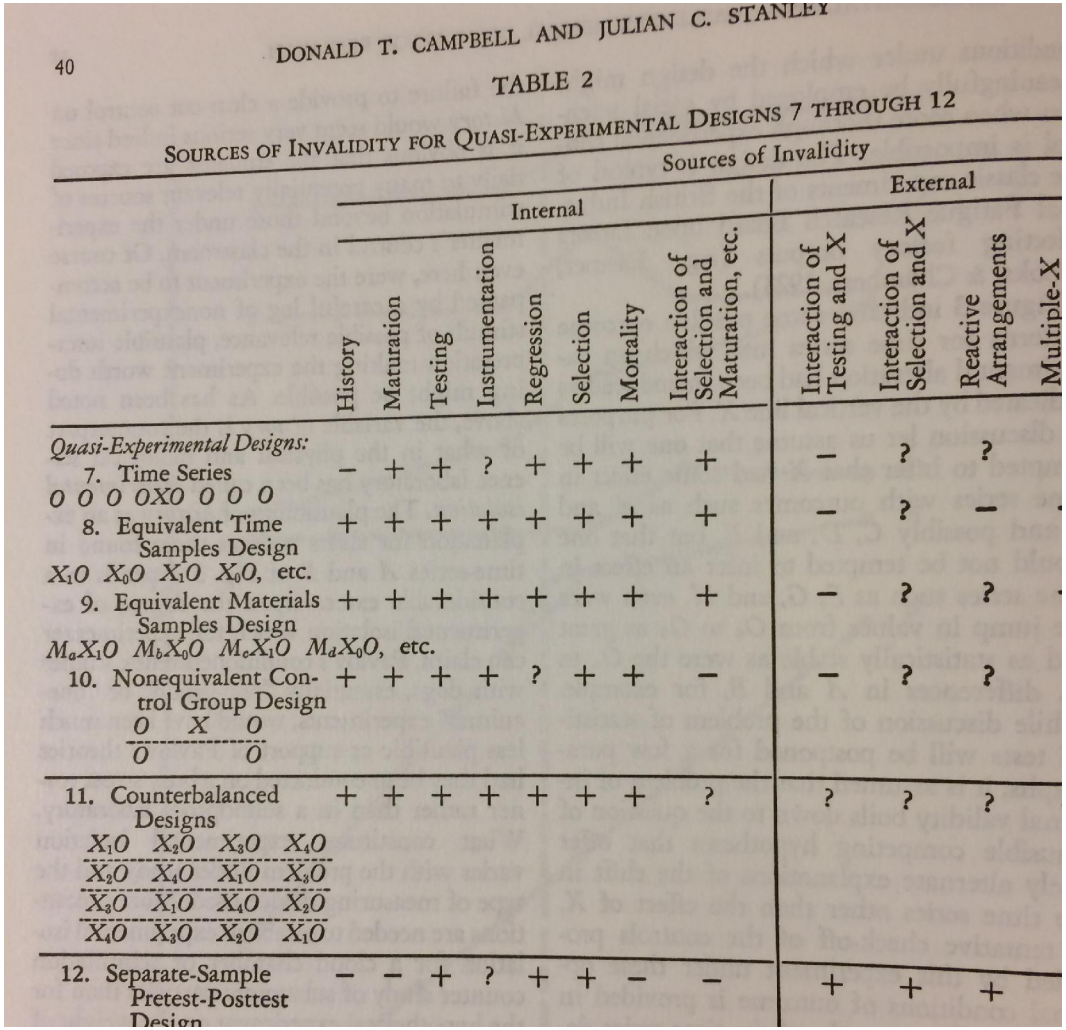


TABLE 2

SOURCES OF INVALIDITY FOR QUASI-EXPERIMENTAL DESIGNS 7 THROUGH 12

	Sources of Invalidity											
	Internal								External			
	History	Maturation	Testing	Instrumentation	Regression	Selection	Mortality	Interaction of Selection and Maturation, etc.	Interaction of Testing and X	Interaction of Selection and X	Reactive Arrangements	Multiple-X
Quasi-Experimental Designs:												
7. Time Series *O O O OXO O O O*	−	+	+	?	+	+	+	+	−	?	?	
8. Equivalent Time Samples Design X_1O X_0O X_1O X_0O, etc.	+	+	+	+	+	+	+	+	−	?	−	
9. Equivalent Materials Samples Design M_aX_1O M_bX_0O M_cX_1O M_dX_0O, etc.	+	+	+	+	+	+	+	+	−	?	?	
10. Nonequivalent Control Group Design $O \quad X \quad O$ $O \qquad O$	+	+	+	+	?	+	+	−	−	?	?	
11. Counterbalanced Designs $X_1O \quad X_2O \quad X_3O \quad X_4O$ $X_2O \quad X_4O \quad X_1O \quad X_3O$ $X_3O \quad X_1O \quad X_4O \quad X_2O$ $X_4O \quad X_3O \quad X_2O \quad X_1O$	+	+	+	+	+	+	+	?	?	?	?	
12. Separate-Sample Pretest-Posttest Design	−	−	+	?	+	+	−	−	+	+	+	

What social scientists can do…

- More experiments
- Calculate effect sizes
- More tools and options

What utility program managers can do…

- Do not succumb to the feeling of hopelessness…
- Talk to social scientists, we operationalize and quantify abstract constructs for a living
  - Translation: “We like the squishy behavioral stuff!”


Questions?
Carol Yin cyin@yinsight.net
Edwin.Hornquist@sce.com
© 2007 Carol Yin