Outcome Evaluation: When to do it and how to get started

Ronda Zakocs, PhD, MPH
Children's Safety Network
May 23, 2007

Ronda Zakocs

Today's agenda

What do we learn from outcome evaluations?

Whether and when should programs be subjected to outcome evaluation?

What are the first steps in planning an outcome evaluation?

Got Outcomes?

Did the program work?

Were the program objectives achieved?

Demand for Outcomes:
Accountability

Are scarce resources being used most efficiently and effectively?

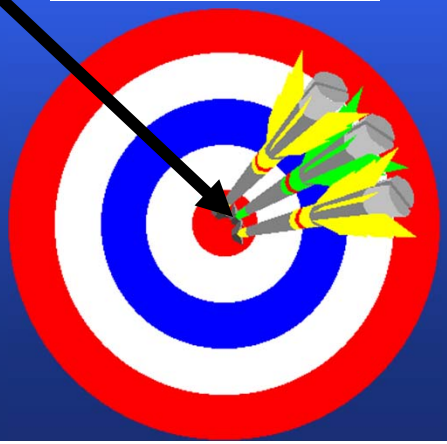
Outcome evaluation seeks to determine whether program was …

Effective

Program prevented childhood deaths and injuries

Program Effort

"Hit the mark"



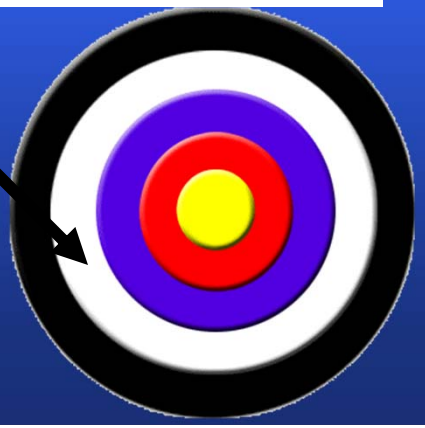
OR

Ineffective

Program did not prevent injuries or deaths or caused adverse effects

Program Effort

Did not "hit the mark"



Outcome evaluation findings on injury prevention programs

Effective Programs

Nurse-Family Partnerships prevents child maltreatment

Widely adopted by many states

Nurse-Family Partnership
Helping First-Time Parents Succeed

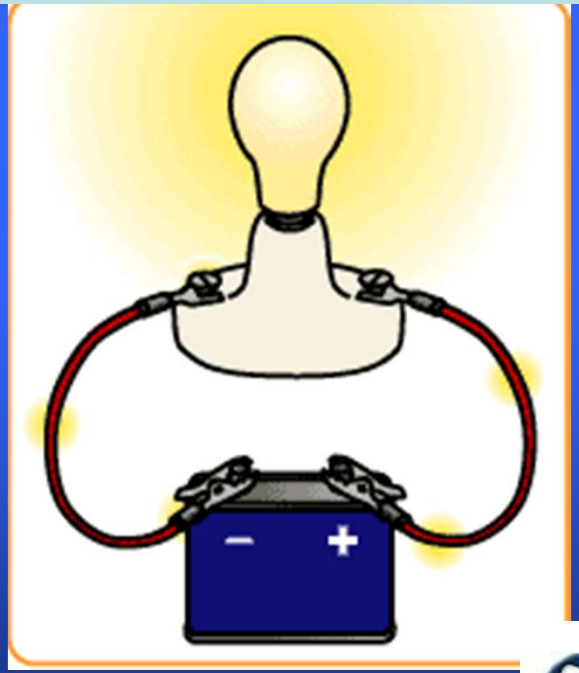
Ineffective Programs

Mr. YUK poison warning stickers do not deter children from touching (and may increase toddler preference) or prevent poisonings

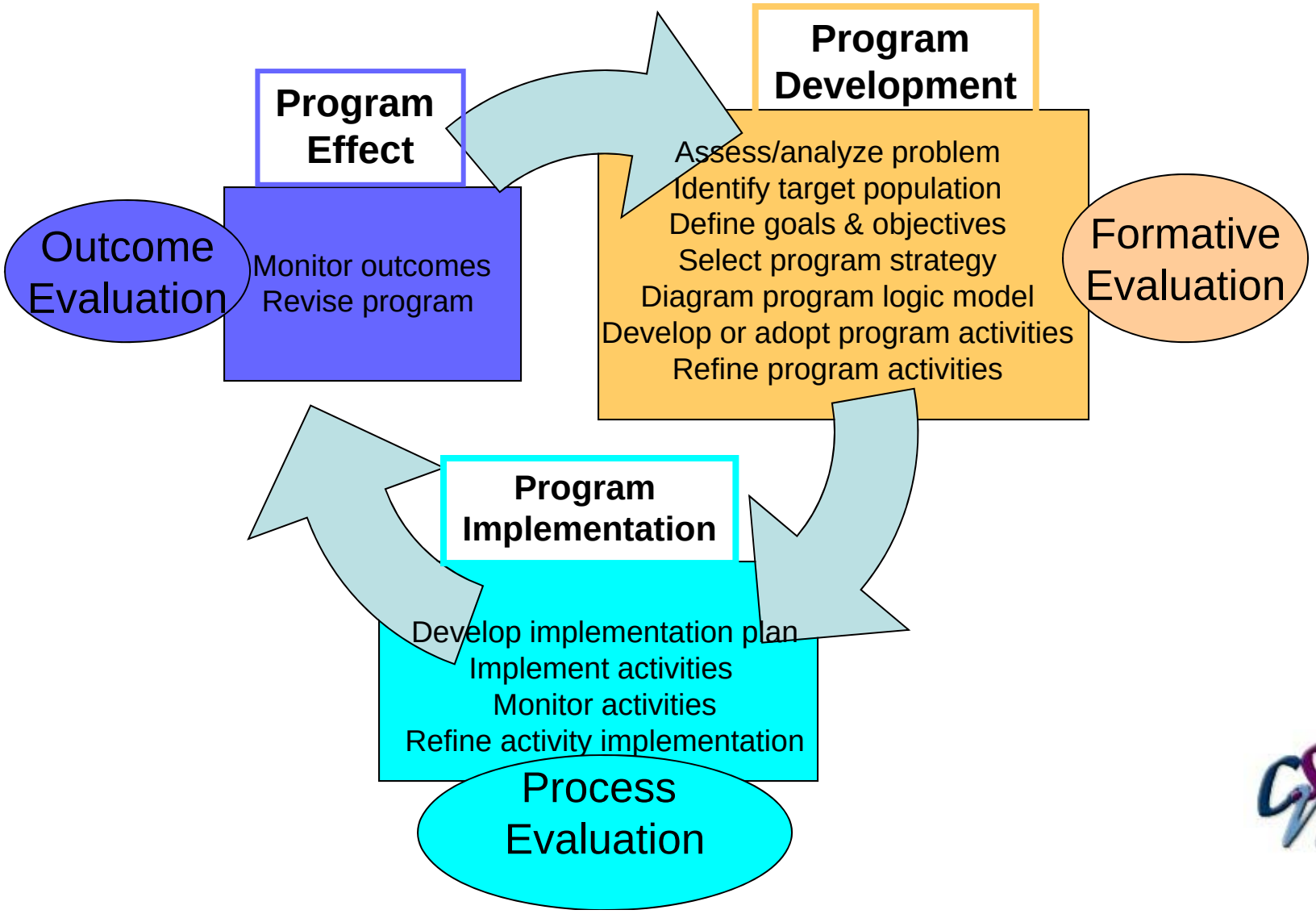
Effective programs tend to be…

Well-designed and conceptually sound

Well-implemented with fidelity



Program lifecycle



Evaluation purposes

Program Development Phase	**Formative Evaluation** Assists in creating well-designed programs
Program Implementation Phase	**Process Evaluation** Assists in implementing well-executed programs
Program Effect Phase	**Outcome Evaluation** Assists in determining program effectiveness

Formative and process evaluation always warranted

Formative evaluation always should be undertaken during program development → Improves better program design

Process evaluation always should be undertaken during program implementation → Improves better program execution

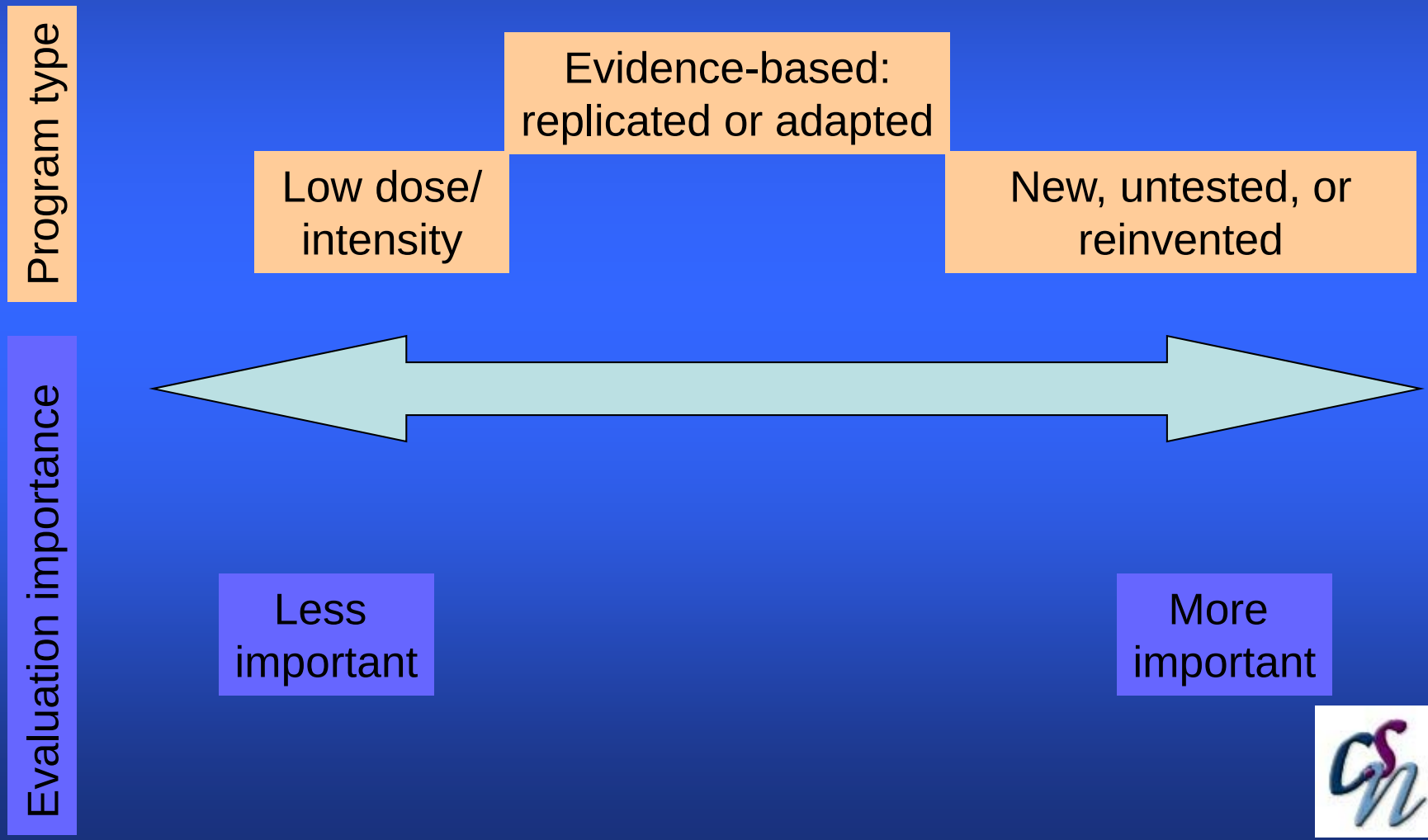
However, outcome evaluation may <u>not</u> be warranted for every program at anytime

Outcome evaluations are resource intensive

- Time
- Money
- Expertise

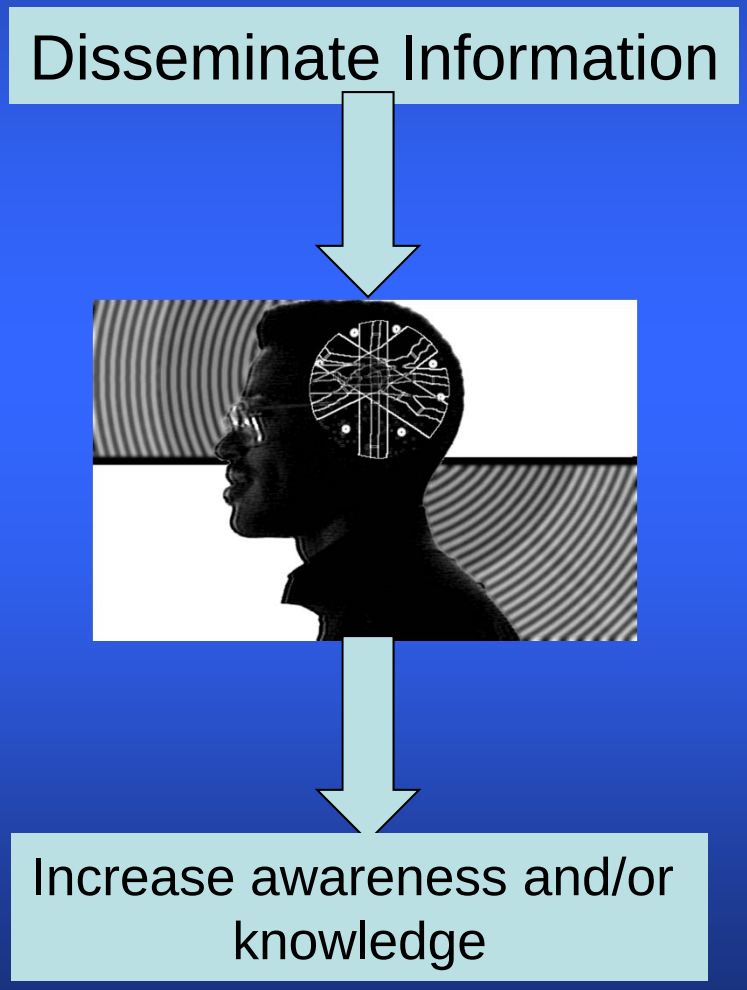
Some programs may be too low dose or already shown to be efficacious

Outcome evaluation decision-making



Low dose/intensity programs

- Disseminate information to increase awareness or knowledge only
- Limited exposure to message
- Examples:
  - Posters
  - Brochures
  - Health fairs
  - Fact sheets
  - Magnets/key chains
  - Information sessions



Outcome evaluations may be less important for low dose programs

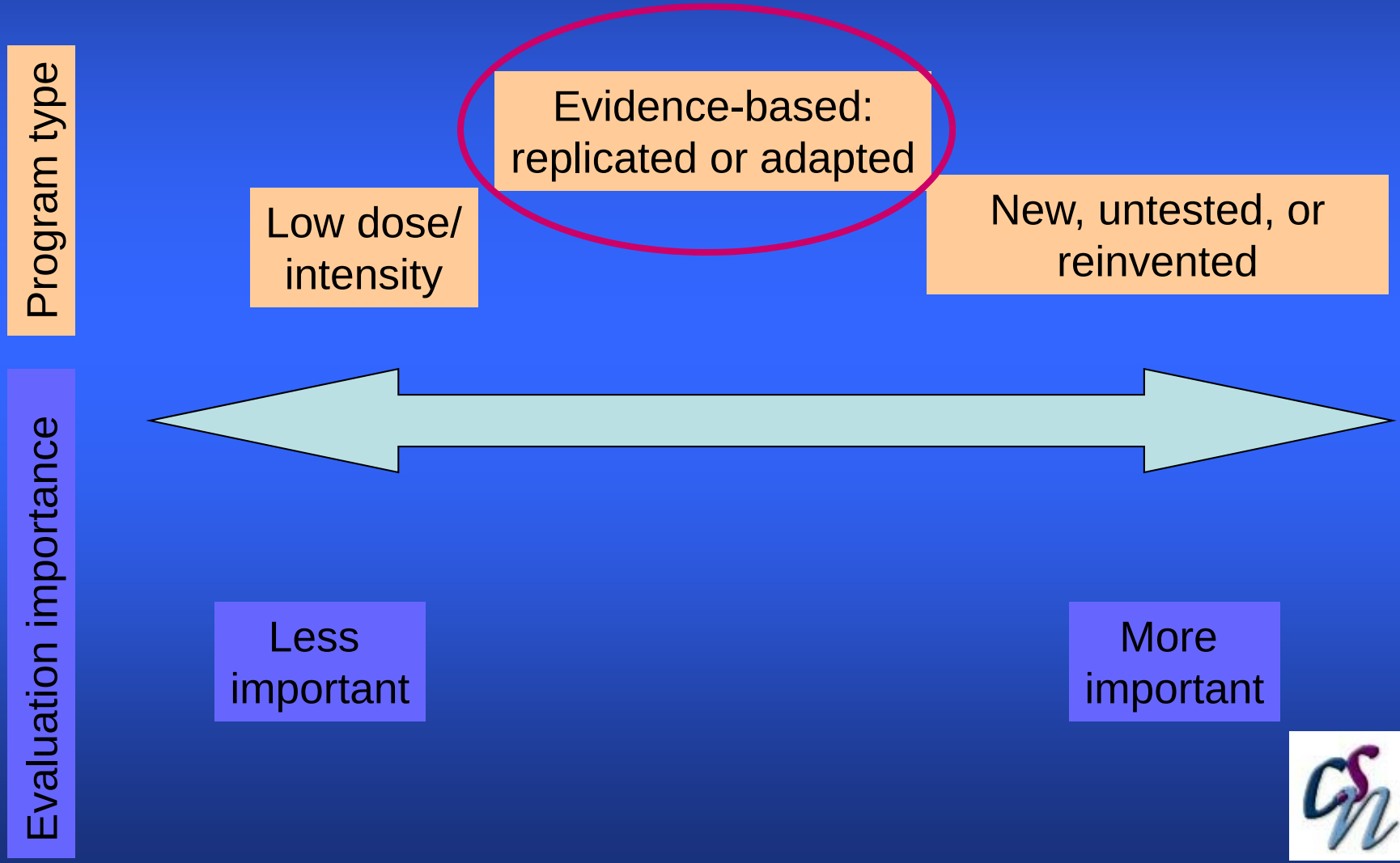
Increasing awareness or knowledge (by itself)

- May be an important foundational activity or one activity of a larger strategy
- BUT research indicates it does not change behavior or socio-environmental conditions ALONE

Therefore….

- Are the resources required for an outcome evaluation justified?
- Might it be more prudent to save scarce resources for other programs or evaluations?

Outcome evaluation decision-making



Evidence-based programs
Definitions

Numerous terms, criteria, and evidence used to identify "evidence-based" programs

- Best practices
- Model program
- Effective program
- Science-based
- Promising program
- Guidelines

For today's webinar, evidence-based programs refers to…

Rigorous, scientific outcome evaluations have determined a program to be effective and, thus, recommended for widespread adoption

Evidence-based programs
Registries and databases

Cochrane Health Promotion and Public Health Collection
- http://cochrane.org
- Reviews of injury prevention housed at:
- http://depts.washington.edu/hiprc/practices/index.html

US PHS Guide to Community Preventive Services
- http://thecommunityguide.org

Best Practices Registry for Suicide Prevention
- http://www.sprc.org/featured_resources/bpr/index.asp

Best practices for youth violence prevention
- http://www.cdc.gov/ncipc/dvp/bestpractices.htm

A Highway Safety Countermeasures Guide for State Highway Safety Offices
- http://www.nhtsa.dot.gov/people/injury/airbags/countermeasures.index.htm

If a program has already been proven effective, then is it necessary to conduct an outcome evaluation when adopted?

Depends …

1. Positive results from evidence-based programs tested in “ideal settings” may not be replicated when adopted in “real world” settings*

2. Adoption of evidence-based program varies greatly*

*Glasgow RE, Lichtenstein E, Marcus AC. Why don’t we see more translation of health Promotion research to practice? *AJPH*. 2003;93:1261-1267

Ways evidence-based programs are adopted*

Replication

- Reproducing program with complete fidelity to protocol and delivered to similar population as in efficacy trial

Adaptation

- Tailoring program to meet needs of different populations or delivery channels
- Core elements remain the same

Re-invention

- Adding or removing core elements

*Taken from Collins, C. (2006) *Evaluating interventions that have already been Determined to be efficacious*. CDC/AEA Summer Evaluation Institute

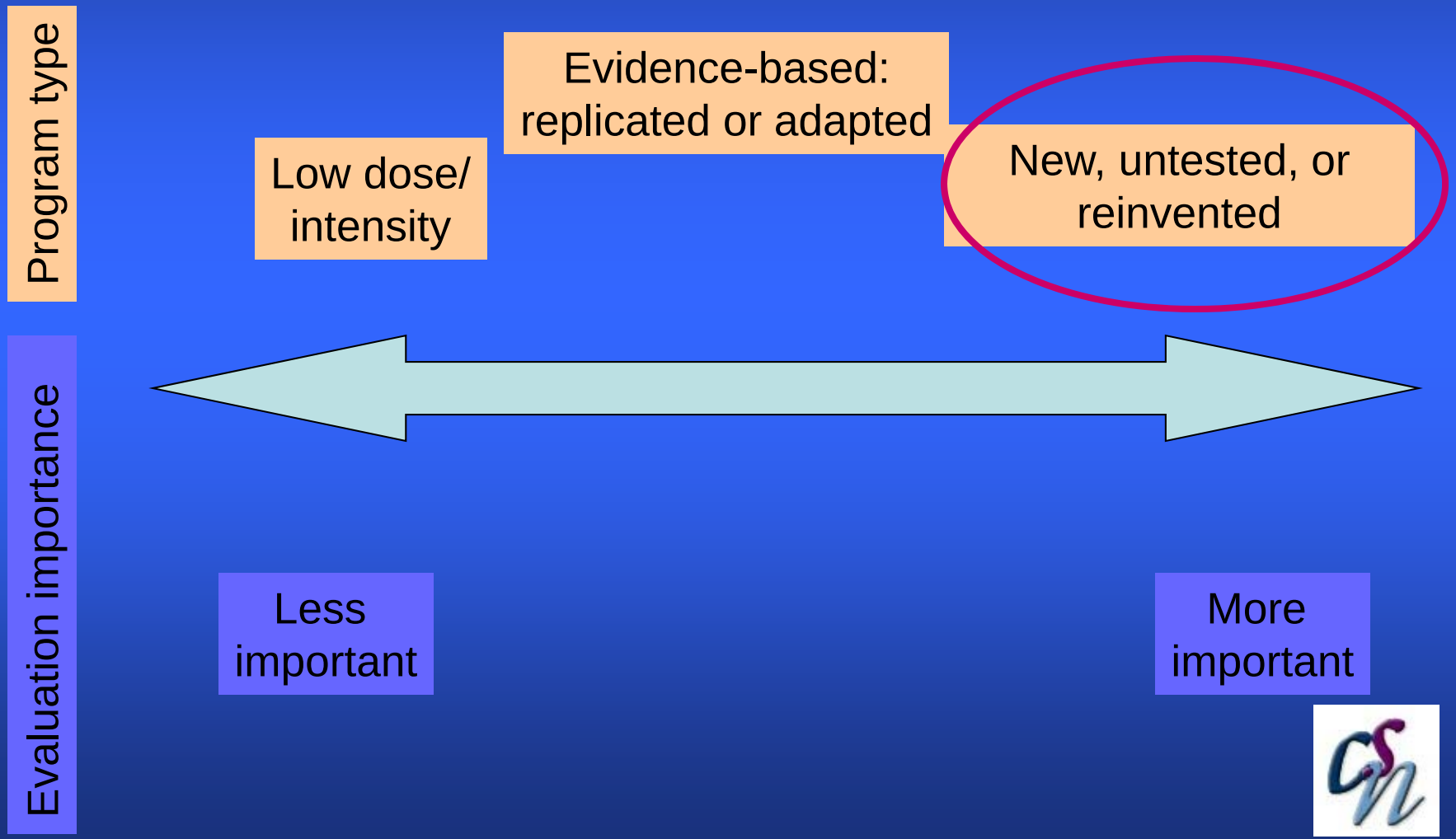
Evidence-based program adoption

*Evaluation requirements**

Evaluation Require-ments	Adoption Type		
	Replication	Adaptation	Re-invention
Formative	X	X	X
Process	X	X	X
Outcome		? depends	X

*Modified from Collins, C. (2006) *Evaluating interventions that have already been determined to be efficacious*. CDC/AEA Summer Evaluation Institute

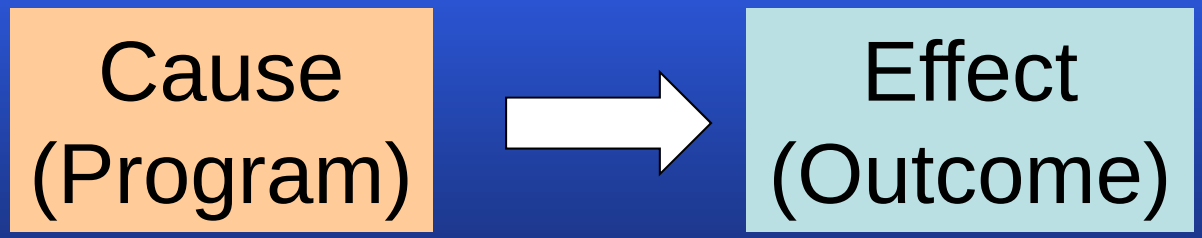
Outcome evaluation decision making



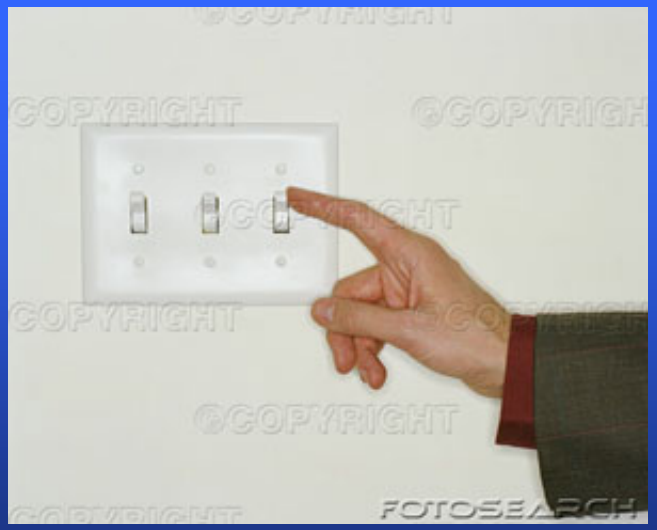
Outcome evaluation always warranted for new, untested, or re-invented programs

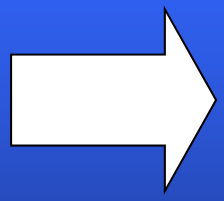
Logic of outcome evaluation

Outcome evaluation is all about attempting to determine whether a program caused an effect in an intended outcome



How do we infer that when a hand "flips a switch" it turns on a light?

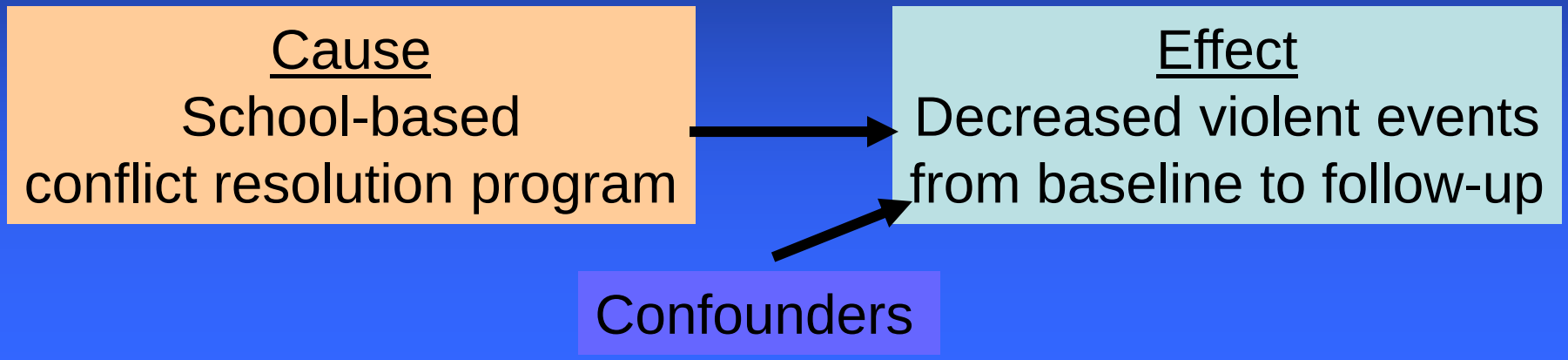




How do we infer a cause and effect relationship?

Criteria #1:	Criteria #2:	Criteria #3:
Is there evidence that the cause preceded the effect?	Is there evidence that the cause changed the effect?	Is there evidence that any other factor (i.e., confounder) caused the effect?
YES: The hand moves before the light turns on.	YES: When the hand moves, the light turns on too.	NO: Do not observe anything else turning on the light. But….

Confounding



What other explanations (i.e., confounders) may explain a decrease in violent events among besides the school-based conflict resolution program?

- New district-wide violence policy
- National violence prevention media campaign
- Community-based, agency coordinated intervention

General outcome evaluation designs

Single group Pre-post test	Quasi-experiment	Randomized control trial
One group: One group receives the program Compares outcome before and after the program	Two groups: One group receives the program and another serves as comparison Compares outcome between program and comparison groups	Two groups: Participants randomly allocated to receive program or serve as control Compares outcome between program and control group

Designs and criteria for inferring causality

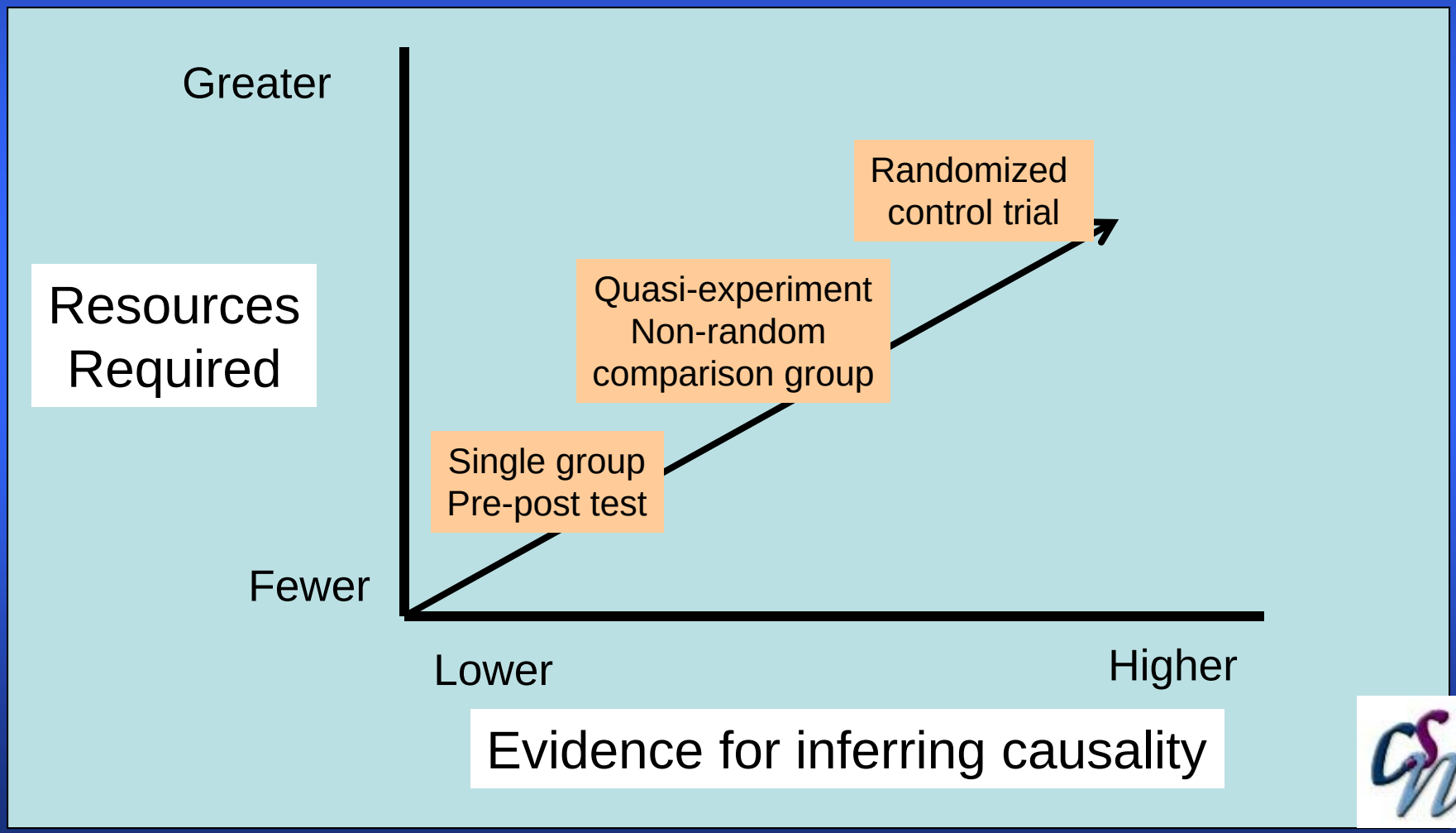
Evidence for inferring causality	Single group, pre-post test	Quasi-experiment	Randomized control trial
(1) Did the program come before outcome?	YES Baseline	YES Baseline	YES Baseline
(2) Did the outcome change in the expected direction?	YES Pre vs post	Stronger Experiment vs control group	Stronger Experiment vs control group
(3) Did something beside the program (e.g., confounder) change outcome?	Weaker No control group	Stronger Non-random comparison group	Strongest Randomization to control group

Why comparison groups provide greater evidence for inferring causality

Helps to rule out alternative explanations (e.g., confounders) for changes found in outcome because groups assumed to be similar except for exposure to program

If positive change found, then comparison groups provide greater evidence that the program was responsible for that change

Stronger evaluation designs tend to require greater resources



What are the first steps in planning an outcome evaluation?

Case Example:

Booster Seat Promotion Program

Booster Seat Promotion Program

Background

- NHTSA recommends booster seats for 4 to 8 year olds
- Boosters reduce injury risk by 59% for 4-7 year olds (Durbin, et al., 2003)
- 72% of CPS seats misused that could increase risk of injury during a crash (NHTSA, 2006)
- 98% infants restrained, yet only 73% of 4-7 year olds restrained (NHTSA, 2005)

Booster Seat Program

- A MCH department will fund and provide TA to community health clinics to deliver the program to low income families over 3 year period
- Low cost booster seat voucher distribution
- Parent education
  - Handouts
  - Nurse counsel

First steps for planning an outcome evaluation

1. Assess readiness
2. Gather needed resources
3. Array possible outcomes by diagramming logic model
4. Select outcomes to evaluate

Assess program readiness

Checklist

- ✓ Is the program well-designed?
  - ➢ *If not, conduct formative evaluation*
- ✓ Is there evidence the program can be implemented as planned?
  - ➢ *If not, conduct process evaluation*
- ✓ Is there a plan on how the results be used?
  - ➢ *If not, then determine how results will be sued*
- ✓ Does your department have the resources?
  - ➢ *Commitment*
  - ➢ *Person power*
  - ➢ *Expertise*

First steps for planning an outcome evaluation

1. Assess readiness
2. Gather needed resources
3. Array possible outcomes by diagramming logic model
4. Select outcomes to evaluate

Gather resources

1. Obtain commitment from higher level administration
2. Assign evaluation coordinator
   - Should not be person responsible for planning/ implementing the program
3. Convene stakeholder evaluation team
   - Should include (at min) higher level administrator, program director, program delivery staff, and evaluation coordinator
4. Gain access to the following expertise either through health department staff, paid or pro bono consultants
   - Evaluation expert with consultation experience
   - Logic model developer
   - Data collector (for whatever outcomes selected)
   - Data programmer/statistician

First steps for planning an outcome evaluation

1. Assess readiness
2. Gather needed resources
3. Array possible outcomes by diagramming logic model
4. Select outcomes to evaluate

Why did Alice get lost?

Alice: Which way should we go?

Cat: That depends on where you are going.

Alice: I don't know where I'm going.

Cat: Then it doesn't matter which way you go!

Lewis Carroll (1872)
Alice in Wonderland

What is a logic model?

- Serves as roadmap for identifying all possible outcomes expected to result from a program
- Diagrams cause (program activities) and effect (expected outcomes) relationships
- Explicitly articulates how a program supposes to work

Logic model elements

Inputs

Resources

Money, staff, or facilities available to implement activities

Strategy/Activities

What the program does with resources aimed at influencing outcomes

Outputs

Indicators of quality and quantity of activities implemented

Outcomes

Short-term

What should change immediately from the activity?

Intermediate

What is *influenced* by the short-term and *influences* the long-term outcome?

Long-term

What is the ultimate injury problem to be addressed by the program?

Source: United Way (1996) *Measuring Program Outcomes*

Outputs vs outcomes

Outputs: Program implementation

Assesses quantity and quality of program activities implemented

Ex: # attending workshop, participant satisfaction

Outcomes: Program effectiveness

Assesses changes in individuals, groups, or environments during or after exposure to program activities

Ex: increase knowledge, reduce injury

Booster Seat Program Logic Model

Inputs

Outcomes

Program Activities

Short-term

Intermediate

Long-term

Low cost booster seat voucher distribution

Parent/caregiver education
• Handouts
•Nurse counsel

Increase access to booster seats

Increase awareness of importance of using booster seats

Increase knowledge/skill on how to use booster seats

Increase booster seat ownership

Increase booster seat use

Increase booster seat correct use

Decrease motor vehicle crash injury rates among 4-8 year olds

Decrease motor vehicle crash death rates among 4-8 year olds

Causes

Effects

What is the difference between impacts and outcomes?

- The literature varies greatly on what is defined as impacts and outcomes

- For today's webinar, prefer to focus on outcomes that display causal linkages among short-term, intermediate, long-term

- But, use whatever terms most comfortable for you

FRIENDS

National Resource Center for Community-Based Child Abuse Prevention

Logic Model Builder

The Logic Model Builder is a collaborative effort between the FRIENDS National Resource Center for Community-Based Child Abuse Prevention, who developed the content and the Child Welfare Information Gateway who developed the database

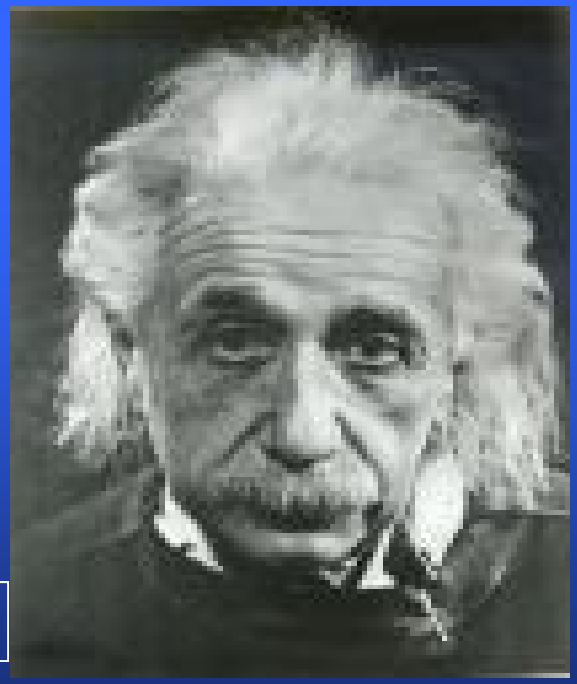
http://www.friendsnrc.org/outcome/toolkit

First steps for planning an outcome evaluation

1. Assess readiness
2. Assemble needed resources
3. Array possible outcomes by diagramming logic model
4. Select outcomes to evaluate

Not all outcomes in logic model must be evaluated; Select outcomes carefully

"Not everything that counts can be counted and not everything that can be counted counts."



Albert Einstein

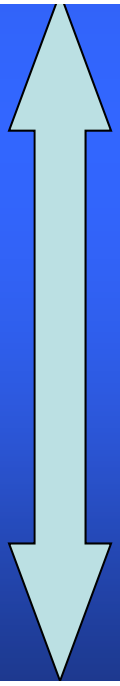
Which outcomes in the logic model…

- Are important to stakeholders?
- Has research already demonstrated causal links?
- Are comparison groups readily available?
- Will there be enough “events” to “rule out chance” for any changes found in outcome?

Which outcomes are important?

Interviewed stakeholders to identify what they wanted to know about the booster seat program

Higher Priority



Were children's lives saved and injuries prevented?

Was there an increase in booster seat use?

Was parental knowledge improved?

Lower Priority

Booster Seat Program Logic Model

Inputs

Program Activities

Low cost booster seat voucher distribution

Parent/caregiver education
• Handouts
•Inspection stations

Outcomes

Short-term

Increase access to booster seats

Increase awareness of importance of using booster seats

Increase knowledge/skill on how to use booster seats

Lower priority

Intermediate

Increase booster seat ownership

Increase booster seat use

Increase booster seat correct use

Moderate priority

Long-term

Decrease motor vehicle crash injury rates among 4-8 year olds

Decrease motor vehicle crash death rates among 4-8 year olds

Higher priority

Has previous research documented causal links among outcomes?

Past epidemiological studies and evaluations may have assessed causal relationships between outcomes in program logic model

If preponderance of literature indicates causal relationship among outcomes in logic model, then may decide not to select those for this evaluation

Booster Seat Program Logic Model

Inputs | **Outcomes**

Program Activities | **Short-term** | **Intermediate** | **Long-term**

May prefer to focus on short and intermediate outcomes

Low cost booster seat voucher distribution

Parent/caregiver education
• Handouts
•Inspection stations

Increase access to booster seats

Increase awareness of importance of using booster seats

Increase knowledge/skill on how to use booster seats

Increase booster seat ownership

Increase booster seat use

Increase booster seat correct use

Decrease motor vehicle crash injury rates among 4-8 year olds

Decrease motor vehicle crash death rates among 4-8 year olds

Red lines indicate substantial epidemiology research has demonstrated causal relationships among use and deaths/injuries.

Are comparison groups available?
Challenges for using comparison groups

1. Identifying and gaining access to appropriate comparison groups
   - Program and comparison groups should be as similar as possible with the only main difference being one received a program and the other did not

2. Costs associated with collecting outcome data for comparison groups
   - Primary data (collect data for evaluation) increases costs
   - Secondary data (analyze existing data) tends to decrease costs

Are comparison groups available?

1. Where could similar families with 4 to 8 year olds not participating in the program be found?
   - Other community health clinics with similar patient demographic profiles?
   - Families residing in the same communities receiving WIC?

2. Are there outcomes for which existing databases could be analyzed at the community level? (Help reduce costs)
   - US Fatal Analysis Reporting System (FARS) to measure deaths?
   - Behavioral Risk Factor Surveillance System (BRSS) measure booster seat use?

Booster Seat Program Logic Model

Inputs | **Outcomes**

Program Activities | **Short-term** | **Intermediate** | **Long-term**

Low cost booster seat voucher distribution

Parent/caregiver education
• Handouts
•Inspection stations

Increase access to booster seats

Increase awareness of importance of using booster seats

Increase knowledge/skill on how to use booster seats

Increase booster seat ownership

Increase booster seat use

Increase booster seat correct use

Decrease motor vehicle crash injury rates a among 4-8 year olds

Decrease motor vehicle crash death rates a among 4-8 year olds

Most likely no databases exist where any of these outcomes could be measured. Would need to collect outcome data for both program and comparison group.

Will there be enough events to rule out chance for any change in outcomes?

Background

- Events
  - Number of data points for measuring an outcome – injury or death rates
- Rule out chance
  - Typically, statistical tests used to assess whether change found between the program and comparison was due to chance – meaning that if another sample had been selected would the same change have been found?
  - Typically, it's easier to rule out chance with larger samples (more events) and greater differences found in outcome between program and comparison groups

Will there be enough events to rule out chance for any change in outcomes? *Example*

- Observe booster seat use at post-program between program group (80%) and comparison group (70%).
- Statistical test used to determine whether 10% difference was due to chance. If yes, then report there was no statistically significant difference between program and comparison group.

Booster Seat Program Logic Model

Inputs

Program Activities

Low cost booster seat voucher distribution

Parent/caregiver education
• Handouts
•Inspection stations

Outcomes

Short-term | **Intermediate** | **Long-term**

More events; greater likelihood of statistical significance

Increase access to booster seats

Increase awareness of importance of using booster seats

Increase knowledge/skill on how to use booster seats

Increase booster seat ownership

Increase booster seat use

Increase booster seat correct use

Too few events; small changes may not be statistically significant

Decrease motor vehicle crash injury rates a among 4-8 year olds

Decrease motor vehicle crash death rates a among 4-8 year olds

Which outcomes to select?

May decide to rule out

- Deaths and injuries
  - Causal evidence already links booster seat use and death/injury reductions
  - Too few events; small changes may not be statistically significant
- Awareness/knowledge
  - Low priority for stakeholders
  - Knowledge alone does not typically influence behavior

May consider selecting

- Ownership
  - Unknown whether a program can successfully increase ownership and whether ownership leads to use
- Use
  - Moderate priority for stakeholders
  - Key causal link for reducing injuries and deaths

Booster Seat Program Logic Model

Inputs

Program Activities

Low cost booster seat voucher distribution

Parent/caregiver education
• Handouts
•Inspection stations

Outcomes

Short-term

Increase access to booster seats

Increase awareness of importance of using booster seats

Increase knowledge/skill on how to use booster seats

Intermediate

Outcomes to Evaluate

Increase booster seat ownership

Increase booster seat use

Increase booster seat correct use

Long-term

Decrease motor vehicle crash injury rates a among 4-8 year olds

Decrease motor vehicle crash death rates a among 4-8 year olds

Evaluation Resources

Outcome evaluation "How to Handbooks"

Hatry, H; Houten, TV; Plantz, MC; and Taylor, M. (1996) *Measuring program outcomes: A practical approach*. Alexandra, VA: United Way of America.

Thompson NJ; McClintock HO. (2000) *Demonstrating Your Program's Worth: A Primer on Evaluation for Programs to Prevent Unintentional Injury*. Atlanta, GA: Centers for Disease Control and Prevention, National Center for Injury Prevention and Control. http://www.cdc.gov/ncipc/pub-res/demonstr.htm

W.K. Kellogg Foundation (1998) *The W.K. Foundation Evaluation Handbook*. Battle Creek, MI: WK Kellogg Foundation. http://www.wkkf.org/documents/WKKF/EvaluationHandbook/EvalHandbook.pdf

Evaluation Resources

Program Logic Models

FRIENDS Evaluation (electronic). Toolkit contains the Logic Model Builder, which was developed in partnership with the National Clearinghouse on Child Abuse and Neglect Information. http://www.friendsnrc.org/outcome/toolkit

United Way of America (1996) *Measuring Program Outcomes: A Practical Approach*. Alexandria, VA: United Way of America.

W.K. Kellogg Foundation (2003) *Logic Model Development Guide*. Battle Creek, MI.