Summing it up

Practical guidance for public health program planning, evaluation, and data use

Presented by:
Michele Polacsek,
Associate Professor of Public Health
Center for Community and Public Health
University of New England
mpolacsek@une.edu
and
and Liam O'Brien,
Associate Professor of Statistics
Colby College
lobrien@colby.edu

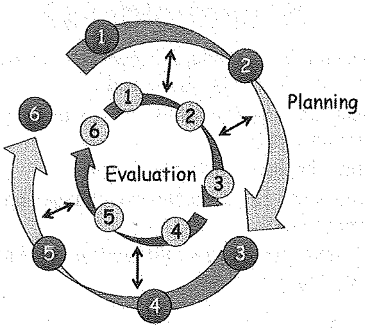
Goals

- Understand how planning and evaluation are connected
- Describe basic steps in public health program Planning
- Describe common steps in public health program evaluation
- Feel comfortable creating a simple overview logic model
- Understand how to write goals and objectives

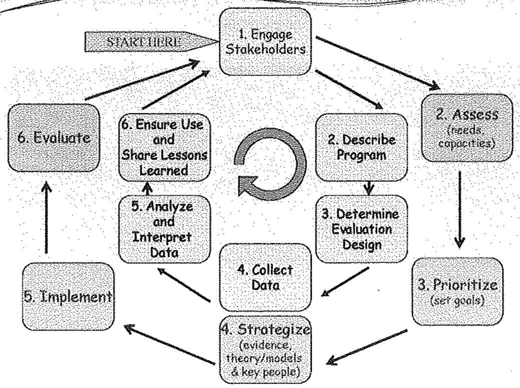
Goals

- Understand the difference between quantitative and qualitative data
- Identify different evaluation designs as well as their benefits and limitations
- Understand the limitations of common sampling strategies
- Describe the properties of a useful survey
- Know what hypotheses are and how they are used
- Be aware of typical challenges in interpreting and analyzing data

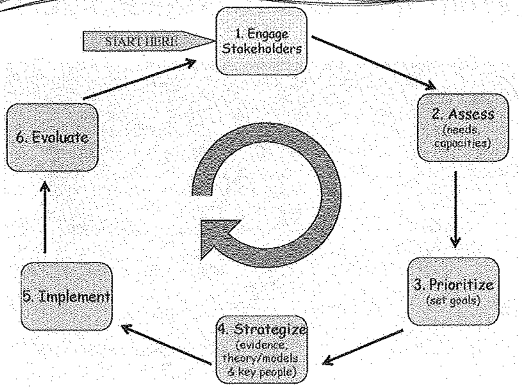
Planning and Evaluation Cycles



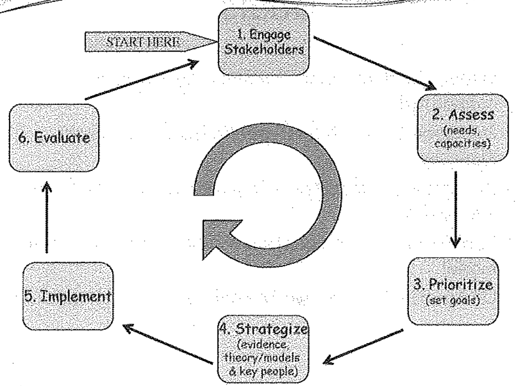
Program Planning and Evaluation Overview


START HERE
1. Engage Stakeholders
2. Assess (needs, capacities)
3. Prioritize (set goals)
4. Strategize (evidence, theory/models & key people)
5. Implement
6. Evaluate
2. Describe Program
3. Determine Evaluation Design
4. Collect Data
5. Analyze and Interpret Data
6. Ensure Use and Share Lessons Learned

Public Health Program Planning


START HERE
1. Engage Stakeholders
2. Assess (needs, capacities)
3. Prioritize (set goals)
4. Strategize (evidence, theory/models & key people)
5. Implement
6. Evaluate

Public Health Program Planning


START HERE
1. Engage Stakeholders
2. Assess (needs, capacities)
3. Prioritize (set goals)
4. Strategize (evidence, theory/models & key people)
5. Implement
6. Evaluate

Stakeholders?

People or organizations invested in the program, who have an interest in the results, and/or have a stake in what will be done with the results

- *People involved in program operations*
  (managers, staff, funders, coalition members)
- *Those served or affected*
  (patients, clients, advocacy groups, community members, elected officials)
- *Users*
  (policy makers, funders, taxpayers, general public, program critics)

Stakeholders?

- Help or hinder planning and evaluation *before*, *while*, and *after* it is conducted
- Can increase the credibility of the program and evaluation results
- May help implementation efforts
- May advocate for or authorize changes based on program results
- May fund or authorize continuation or expansion of the program

Identifying Stakeholders

(Example: Childhood Lead Poisoning Prevention)

Who are the key stakeholders we need to:			
Increase credibility of our efforts	Implement the interventions that are central to this effort	Advocate for changes to institutionalize this effort	Fund/authorize continuation or expansion of this effort
Physician associations Community associations	State and local health departments Housing authorities	Advocacy groups Maternal and child health groups Physician associations Community associations	Legislators and policymakers at federal and state levels (M)CDC Private Industry Court system

What Matters to Stakeholders

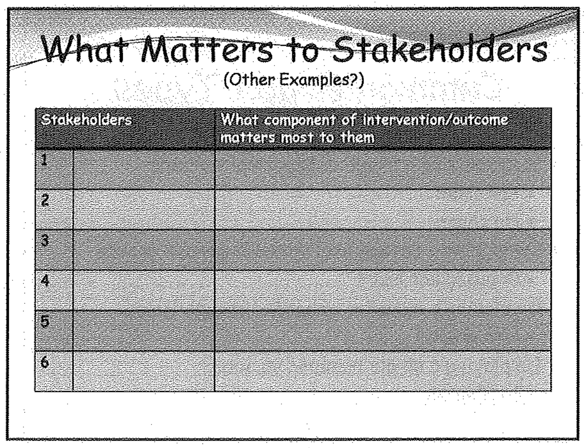
(CLPP)

Stakeholders		What component of intervention/outcome matters most to them
1	Physician associations	Sufficient "yield" of EBLL children to make their screening efforts worth their time Clear referral mechanisms that are easy and work
2	Community associations	Cleaning up housing in their neighborhood Support for families with EBLL children
3	Housing authorities	No additional monetary and time burden for toxic cleanups
4	State/local health departments	Efforts lead to improved health outcomes for EBLL children
5	Advocacy groups	EBLL is seen as a housing problem and not a "parenting" failure
6	Policymakers	Efforts lead to improved health outcomes "cost-effectiveness" of the effort

Identifying Stakeholders

(Other Examples?)

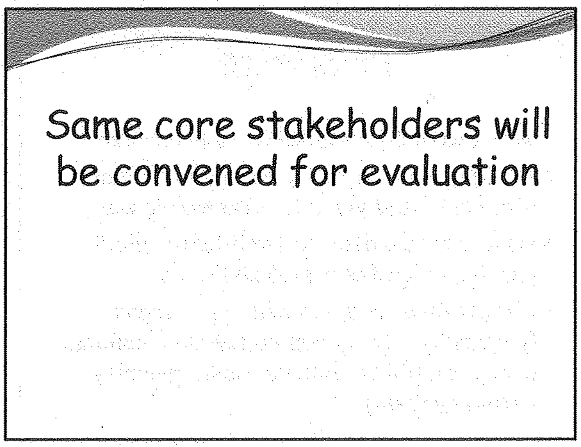
Who are the key stakeholders we need to: *Remember: operations, those affected and users*			
Increase credibility of our efforts	Implement the interventions that are central to this effort	Advocate for changes to institutionalize this effort	Fund/authorize continuation or expansion of this effort



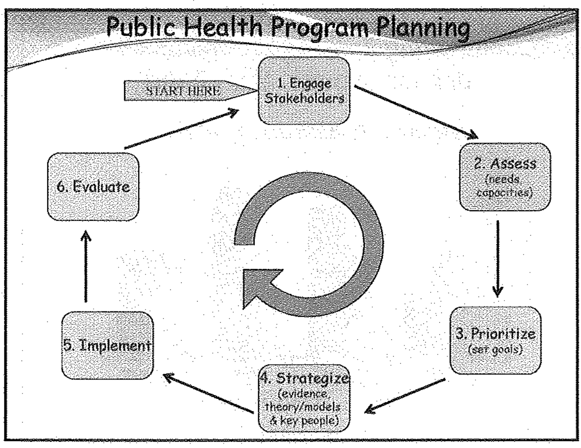
What Matters to Stakeholders

(Other Examples?)

	Stakeholders	What component of intervention/outcome matters most to them
1		
2		
3		
4		
5		
6		



Same core stakeholders will be convened for evaluation



Public Health Program Planning

START HERE

1. Engage Stakeholders

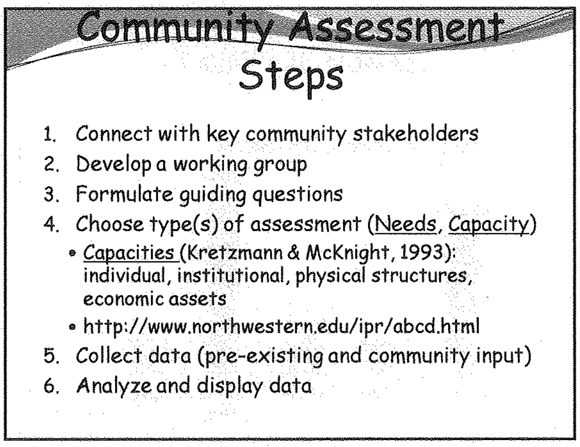
2. Assess (needs, capacities)

3. Prioritize (set goals)

4. Strategize (evidence, theory/models & key people)

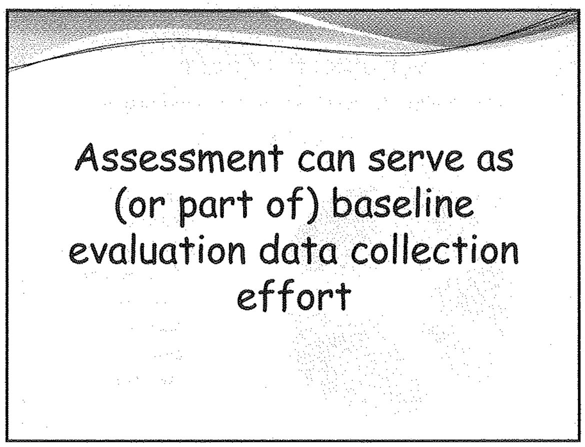
5. Implement

6. Evaluate

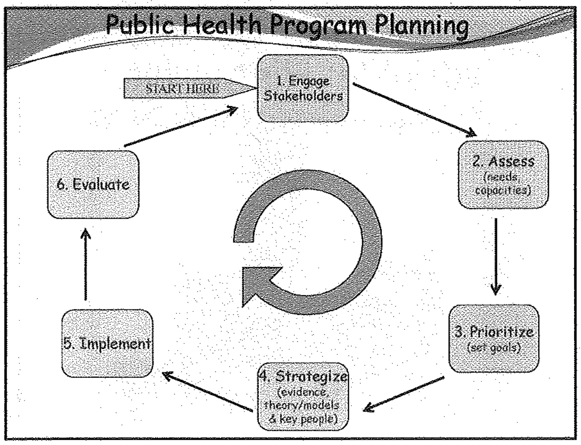


Community Assessment Steps

1. Connect with key community stakeholders
2. Develop a working group
3. Formulate guiding questions
4. Choose type(s) of assessment (Needs, Capacity)
    - Capacities (Kretzmann & McKnight, 1993): individual, institutional, physical structures, economic assets
    - http://www.northwestern.edu/ipr/abcd.html
5. Collect data (pre-existing and community input)
6. Analyze and display data



Assessment can serve as (or part of) baseline evaluation data collection effort



Public Health Program Planning

START HERE

1. Engage Stakeholders

2. Assess (needs, capacities)

3. Prioritize (set goals)

4. Strategize (evidence, theory/models & key people)

5. Implement

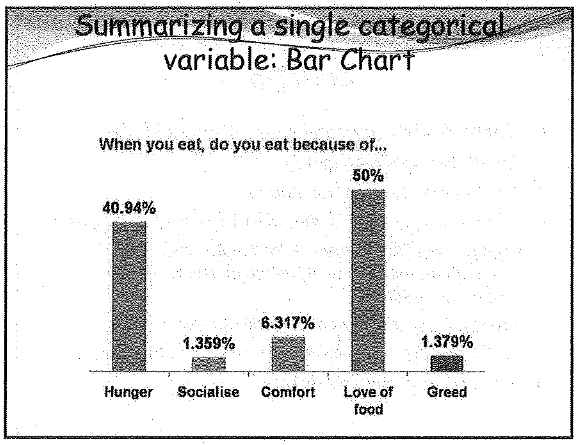
6. Evaluate

Prioritize

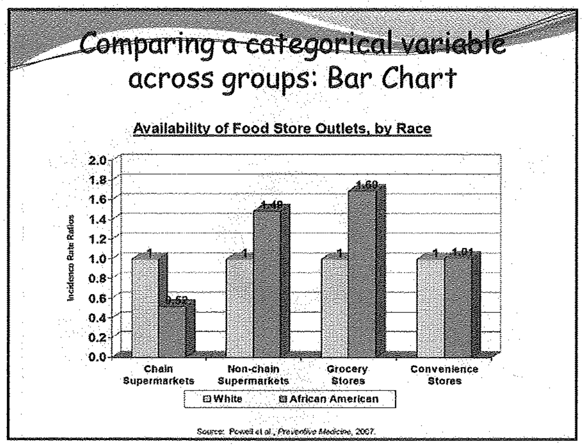
- Work with Stakeholders to prioritize
- Present assessment findings in a clear, organized and visually interesting way
- Have a moderator or facilitator elicit participation from stakeholders
- Decide on a decision making process (majority vote, group consensus, nominal group, delphi technique, basic priority rating system)

Common Graph Types

- Summarizing a single categorical variable
  - Bar chart
  - Pie chart
- Comparing a categorical variable across groups
  - Grouped bar chart
- Displaying a numeric variable through time
  - Line graph
- Displaying a single numeric variable
  - Histogram

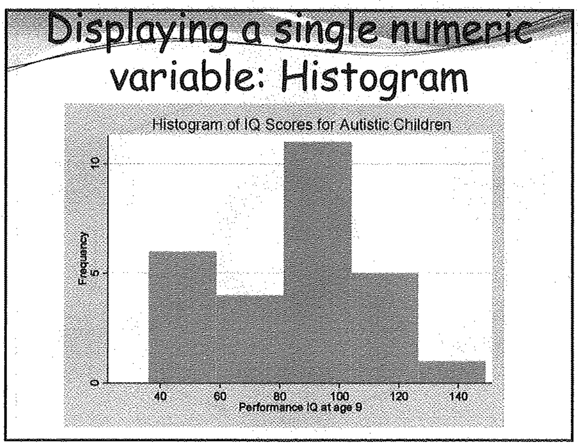


Summarizing a single categorical variable: Bar Chart

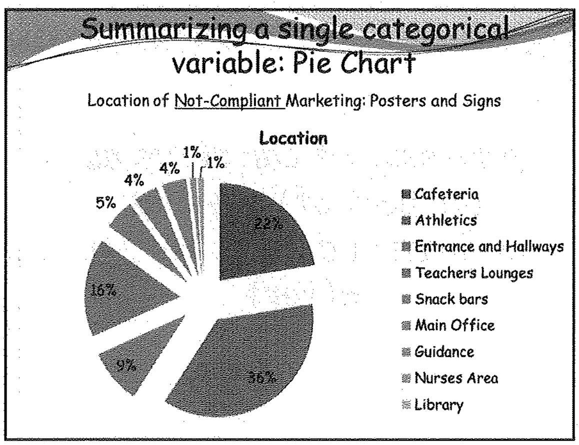


Comparing a categorical variable across groups: Bar Chart

Source: Powell et al., *Preventive Medicine*, 2007.



Displaying a single numeric variable: Histogram

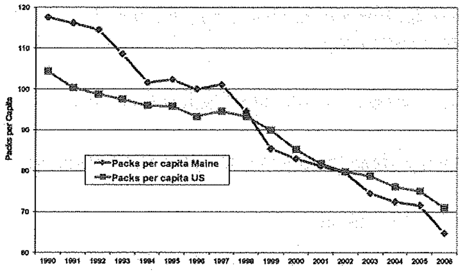


Summarizing a single categorical variable: Pie Chart

Location of Not-Compliant Marketing: Posters and Signs

Displaying a numeric variable through time: Line graph

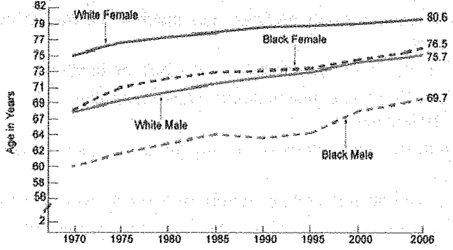
Cigarette Consumption - Packs Sold Per Capita
Maine and US 1990-2006



Source: The Burden on Tobacco, Orzechowski and Walker

Comparing a numeric variable through time across groups: Line graph

Life Expectancy at Birth, by Race* and Sex, 1970–2006
Source II.4: Centers for Disease Control and Prevention, National Center for Health Statistics



**Both racial categories include Hispanics.*

Principles of Effective Graphs

- Show the data clearly
- Represent magnitudes accurately
- Minimize clutter
- Make displays easy to interpret
- Clearly identify the axes

Same lessons for presenting evaluation data

Prioritize

- THEN using assessment data: stakeholders prioritize:
  - existing needs/gaps
  - resources/capacities
  - barriers to using existing resources
  - where new intervention is warranted

Techniques to help a group prioritize

- nominal group
- delphi technique
- basic priority rating system

Nominal Group Technique

The **nominal group technique** (NGT) is a decision-making method for use among group of varying sizes

- When some group members are much more vocal than others.
- When some group members think better in silence.
- When there is concern about some members not participating.
- When the group does not easily generate quantities of ideas.
- When all or some group members are new to the team.
- When the issue is controversial or there is heated conflict.

Steps in the NGT

1. Introduction and explanation (sharing of assessment data)
2. Silent generation of ideas/issues (priorities)
3. Sharing ideas (everyone: one at a time sharing)
4. Group discussion
5. Anonymous voting and finally ranking issues

Delphi Technique

- The Delphi method is an iterative structured communication technique used to elicit common judgment(s) from a group of experts.
  - experts answer questionnaires in two or more rounds (stakeholders prioritize with rationale via questionnaire)
  - facilitator provides summary of judgments
  - experts revise their earlier answers in light of the replies of other members
  - the range of the answers will decrease and the group will converge
  - Mean or median scores of the final rounds determine the results

Basic Priority Rating System

(Like Importance/Changeability matrix)

Asks group members to rank assessment findings based on 3 components:

1. The size of the problem
2. The seriousness of the problem
3. The estimated effectiveness of the solution (changeability of the problem)

The highest ranked problem receives priority

Basic Priority Rating System

Practice Exercise

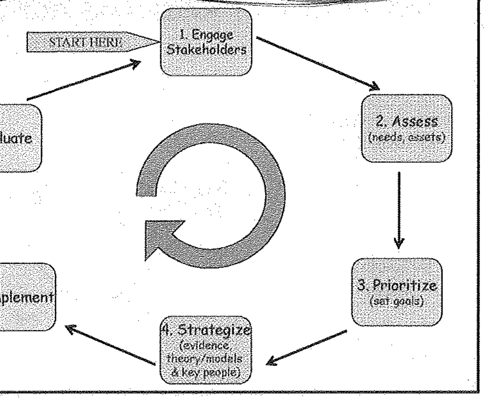
- Unhealthy school vending
  - (e.g. offer more healthy choices in vending)
- Lack of healthy choices for school lunch
  - (e.g. offer salad bar every day)
- Marketing of unhealthy foods at school
  - (e.g. remove all food and beverage marketing)
- Students bring unhealthy lunches to school
  - (e.g. educational campaign to encourage students to bring healthier lunches)

Set Goal

- State what you want to accomplish in broad terms (and include the population)
  - e.g. Reduce the number of pregnancies in Middletown High School district
  - e.g. Prevent falls among residents of the Center Senior apartments
  - e.g. Improve the cardiovascular health of African American women in Smithville

The goal can become an initial, intermediate, or long-term evaluation objective

Public Health Program Planning


START HERE
1. Engage Stakeholders
2. Assess (needs, assets)
3. Prioritize (set goals)
4. Strategize (evidence, theory/models & key people)
5. Implement
6. Evaluate

Choose Strategies

- **Examine the evidence**
- **Examine theory/models** (e.g. ecological model)
- **Consult with key people/professionals**

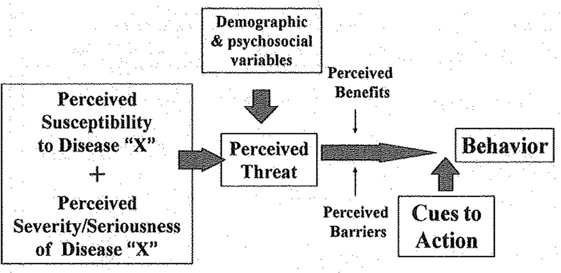
Examine the evidence

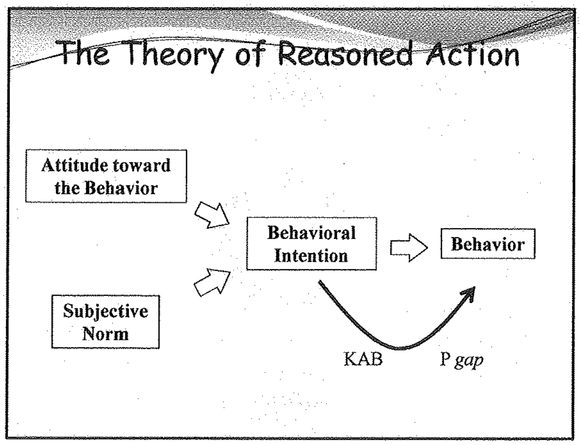
- Google Scholar or other key databases (e.g. PubMed or Web of Science)
- Community Guide (www.thecommunityguide.org)
- Cochrane Review (www.cochrane.org/reviews)
- Health-Evidence.ca (http://health-evidemce.ca)
- National Guideline Clearinghouse (www.guidelines.gov)

Theories/Models

- Health Belief Model
- Theory of Reasoned Action (Planned Behavior)
- Social Learning Theory
- Transtheoretical Model (Stages of Change)
- Social Marketing Model
- Diffusion of Innovations
- Preceed/Proceed
- Social Ecological Model

The Health Belief Model


Demographic & psychosocial variables
Perceived Susceptibility to Disease "X"
+
Perceived Severity/Seriousness of Disease "X"
Perceived Threat
Perceived Benefits
Perceived Barriers
Behavior
Cues to Action



The Theory of Reasoned Action

Attitude toward the Behavior

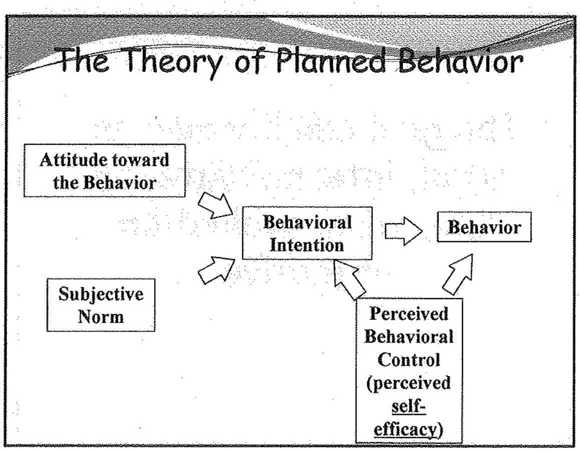
Subjective Norm

Behavioral Intention

Behavior

KAB

P *gap*



The Theory of Planned Behavior

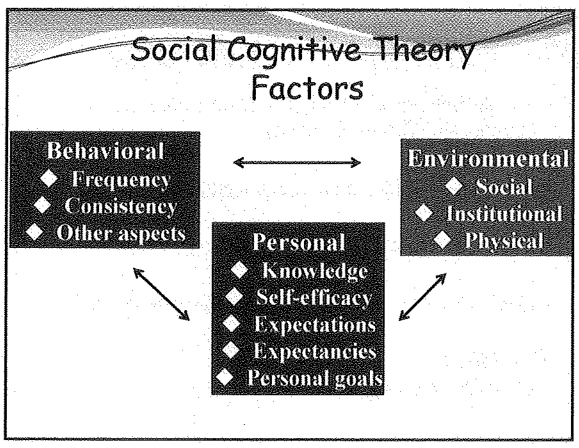
Attitude toward the Behavior

Subjective Norm

Behavioral Intention

Behavior

Perceived Behavioral Control (perceived self-efficacy)

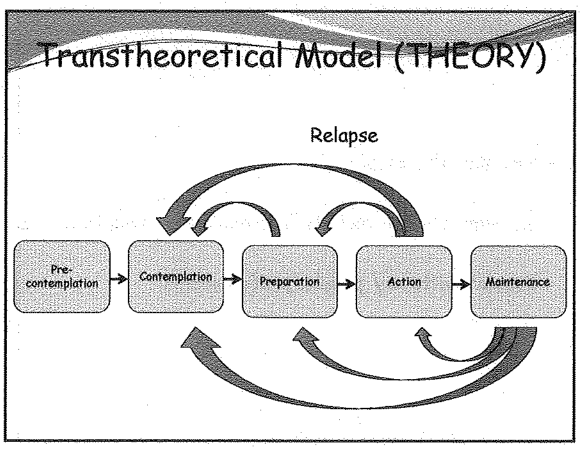


Social Cognitive Theory Factors

Behavioral
- Frequency
- Consistency
- Other aspects

Environmental
- Social
- Institutional
- Physical

Personal
- Knowledge
- Self-efficacy
- Expectations
- Expectancies
- Personal goals



Transtheoretical Model (THEORY)

Relapse

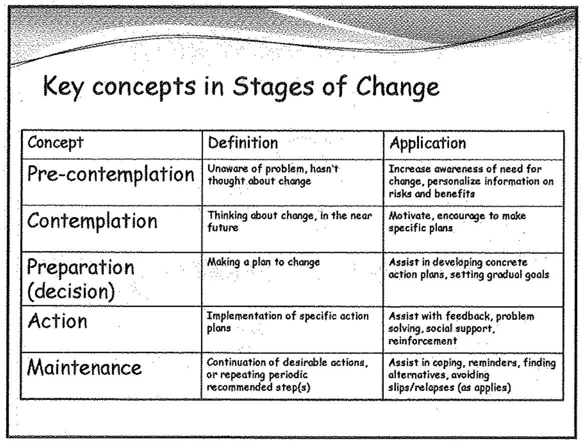
Pre-contemplation

Contemplation

Preparation

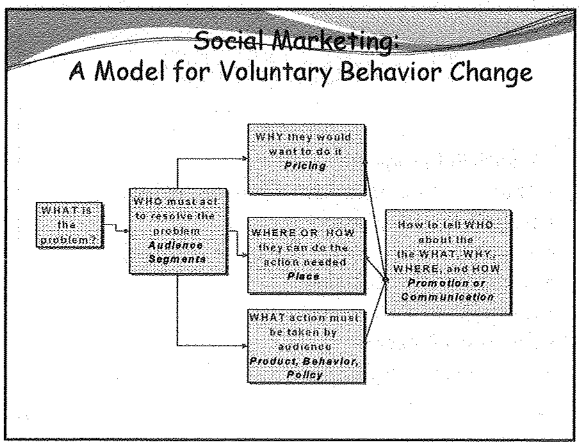
Action

Maintenance



Key concepts in Stages of Change

Concept	Definition	Application
Pre-contemplation	Unaware of problem, hasn't thought about change	Increase awareness of need for change, personalize information on risks and benefits
Contemplation	Thinking about change, in the near future	Motivate, encourage to make specific plans
Preparation (decision)	Making a plan to change	Assist in developing concrete action plans, setting gradual goals
Action	Implementation of specific action plans	Assist with feedback, problem solving, social support, reinforcement
Maintenance	Continuation of desirable actions, or repeating periodic recommended step(s)	Assist in coping, reminders, finding alternatives, avoiding slips/relapses (as applies)



Social Marketing:
A Model for Voluntary Behavior Change

WHAT is the problem?

WHO must act to resolve the problem
Audience Segments

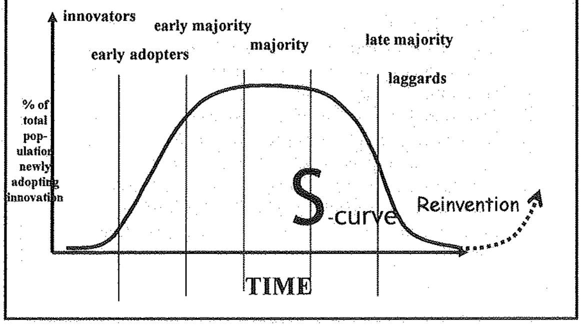
WHY they would want to do it
Pricing

WHERE OR HOW they can do the action needed
Place

WHAT action must be taken by audience
Product, Behavior, Policy

How to tell WHO about the the WHAT, WHY, WHERE, and HOW
Promotion or Communication

Diffusion Of Innovations Curve



Five Maine Constructs in the Adoption/Diffusion of Innovations

- Relative Advantage: innovation better than old behavior
- Compatibility: consistent with existing needs, values, systems
- Complexity: ease of implementation
- Trialability: trial opportunities
- Observability: seeing others do it

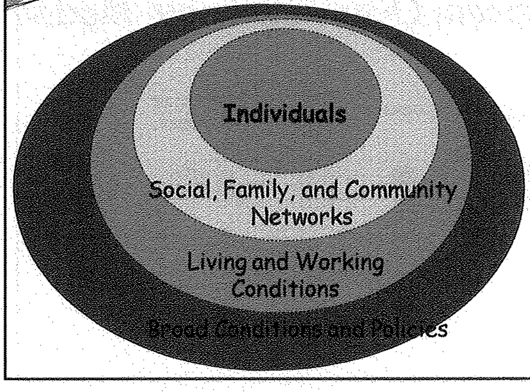
9 Phases of PRECEDE-PROCEED

Diagnostic Phases (5)
- Social (needs, wants, resources, and barriers) -phase 1
- Epidemiological (morbidity, mortality) - phase 2
- Behavioral & Environmental - phase 3
- Educational & Organizational - phase 4
- Administrative & Policy - phase 5

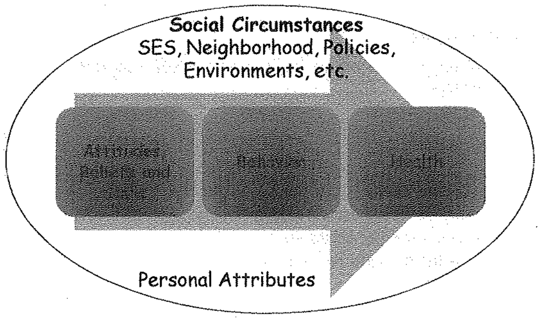
Implementation Phase (1) - phase 6

Evaluation Phases (3)
- Process evaluation - phase 7
- Impact evaluation - phase 8
- Outcome evaluation - phase 9

Social-Ecological Model

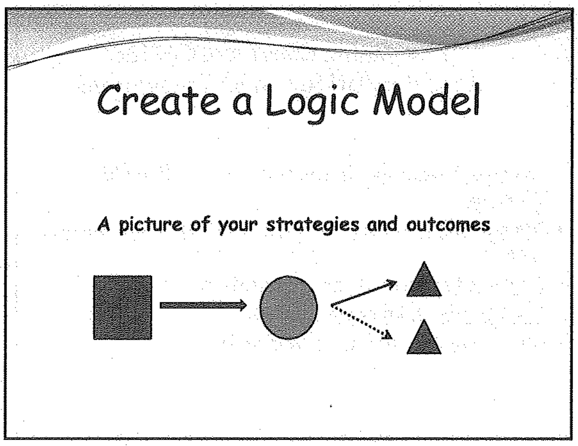

Individuals
Social, Family, and Community Networks
Living and Working Conditions
Broad Conditions and Policies

Summary theory/models


Social Circumstances
SES, Neighborhood, Policies, Environments, etc.

Personal Attributes

Consult with key people/professionals

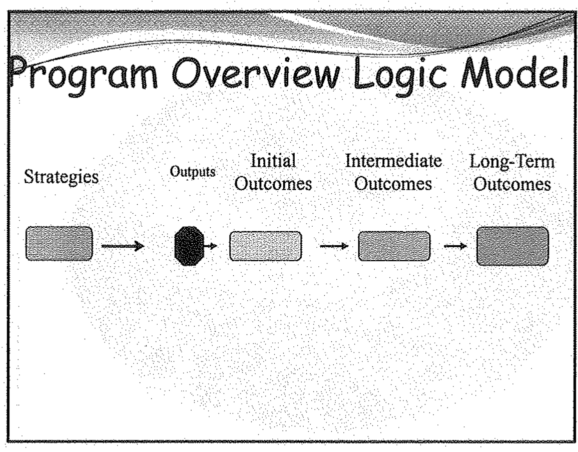


Create a Logic Model

A picture of your strategies and outcomes

Advantages of Logic Models

1. Presents overview i
2. Explains the relevance of a program l
3. Helps programs to plan, set goals o
4. Develops common vision v
5. Creation process fosters understanding e
6. Describes important contextual issues l
7. May reveal unforeseen factors/variables o
8. Strengthens causal claims (program theory) g
9. Can focus on multiple levels of intervention i
10. Creation may involve literature/best practice review c



Program Overview Logic Model

Strategies
Outputs
Initial Outcomes
Intermediate Outcomes
Long-Term Outcomes

Some Definitions

- **Strategies**
  - refers to *doing or accomplishing*
  - the activities you are engaged in
  - examples include meeting, building, equipping, training staff, hiring staff, providing education, providing clinical services, etc.

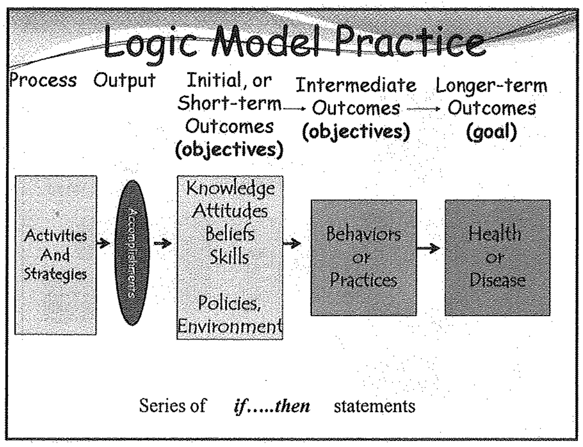
Some Definitions

- **Outputs**
  - Accomplishments or products directly due to the activities and strategies engaged in
  - Outputs are not outcomes. Outputs are the results of activities, accomplishments or products
  - Examples include plans written, meetings held, studies performed, trainings delivered, clinical services provided, or clients served
    - (widgits we count)

Some Definitions

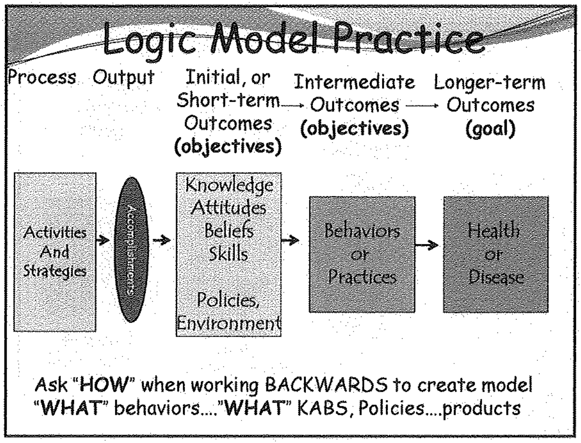
- **Outcomes**
  - Changes that occur ***because of what you are doing***
  - Changes in
    - Individuals' knowledge, attitudes, beliefs, skills (short term)
    - policies, environments (short-term)
    - behaviors (intermediate)
    - health/disease (long-term)

Logic Model Practice


Process
Output
Initial, or Short-term Outcomes **(objectives)**
Intermediate Outcomes **(objectives)**
Longer-term Outcomes **(goal)**
Activities And Strategies
Accomplishments
Knowledge Attitudes Beliefs Skills Policies, Environment
Behaviors or Practices
Health or Disease

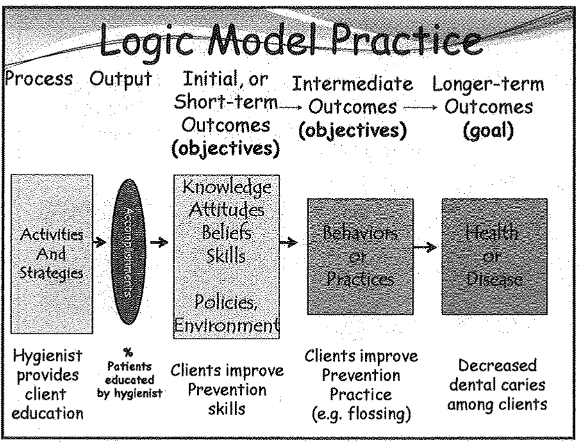
Series of *if.....then* statements

Logic Model Practice

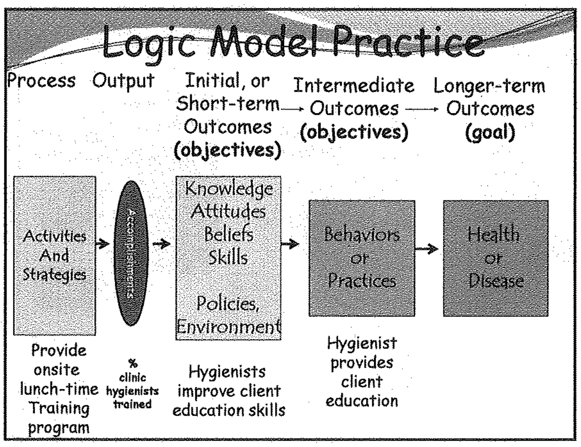

Process
Output
Initial, or Short-term Outcomes **(objectives)**
Intermediate Outcomes **(objectives)**
Longer-term Outcomes **(goal)**
Activities And Strategies
Accomplishments
Knowledge Attitudes Beliefs Skills Policies, Environment
Behaviors or Practices
Health or Disease

Ask "**HOW**" when working BACKWARDS to create model
"**WHAT**" behaviors...."**WHAT**" KABS, Policies....products

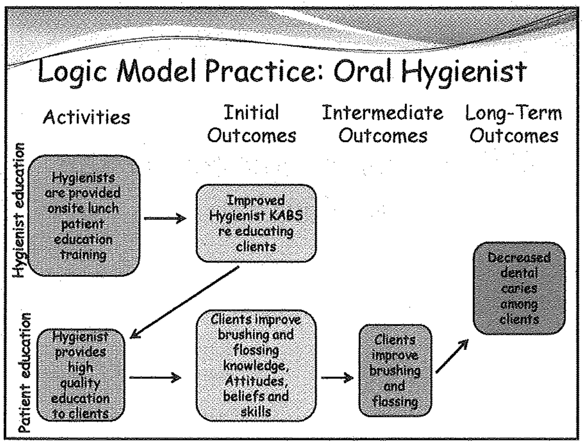
Logic Model Practice


Process
Output
Initial, or Short-term Outcomes **(objectives)**
Intermediate Outcomes **(objectives)**
Longer-term Outcomes **(goal)**
Activities And Strategies
Accomplishments
Knowledge Attitudes Beliefs Skills Policies, Environment
Behaviors or Practices
Health or Disease
Hygienist provides client education
% Patients educated by hygienist
Clients improve Prevention skills
Clients improve Prevention Practice (e.g. flossing)
Decreased dental caries among clients

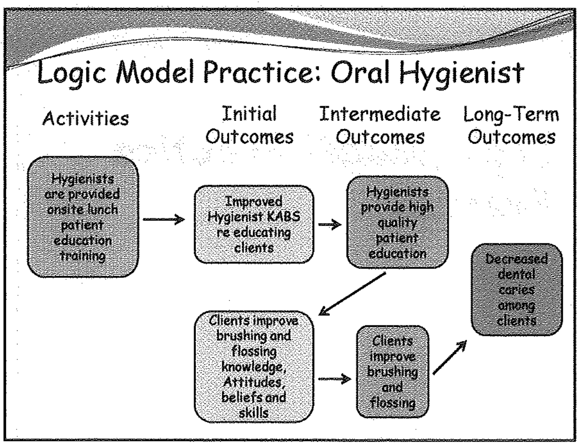
Logic Model Practice

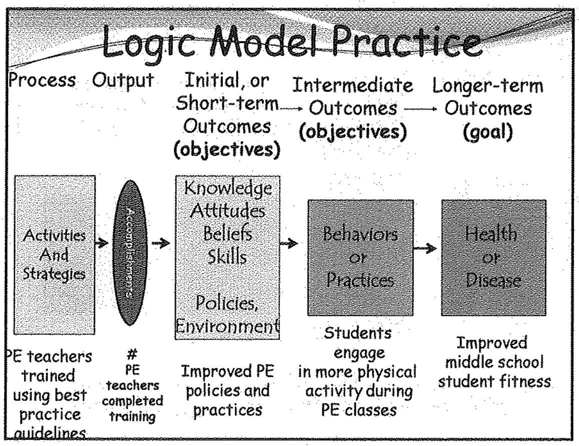

Process
Output
Initial, or Short-term Outcomes **(objectives)**
Intermediate Outcomes **(objectives)**
Longer-term Outcomes **(goal)**
Activities And Strategies
Accomplishments
Knowledge Attitudes Beliefs Skills Policies, Environment
Behaviors or Practices
Health or Disease
Provide onsite lunch-time Training program
% clinic hygienists trained
Hygienists improve client education skills
Hygienist provides client education

Logic Model Practice: Oral Hygienist


Activities
Initial Outcomes
Intermediate Outcomes
Long-Term Outcomes
Hygienist education
Patient education
Hygienists are provided onsite lunch patient education training
Improved Hygienist KABS re educating clients
Hygienist provides high quality education to clients
Clients improve brushing and flossing knowledge, Attitudes, beliefs and skills
Clients improve brushing and flossing
Decreased dental caries among clients

Logic Model Practice: Oral Hygienist


Activities
Initial Outcomes
Intermediate Outcomes
Long-Term Outcomes
Hygienists are provided onsite lunch patient education training
Improved Hygienist KABS re educating clients
Hygienists provide high quality patient education
Clients improve brushing and flossing knowledge, Attitudes, beliefs and skills
Clients improve brushing and flossing
Decreased dental caries among clients

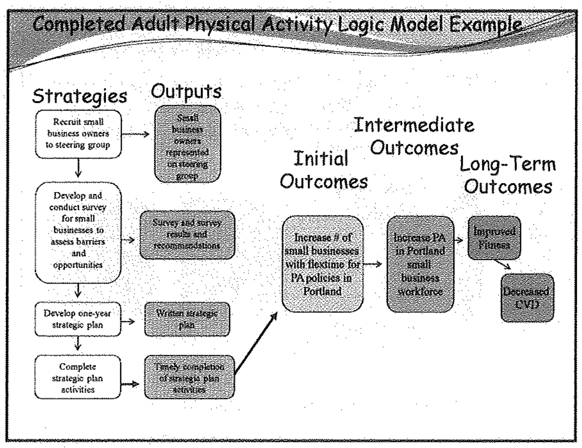


Logic Model Practice

Process
Output
Initial, or Short-term Outcomes **(objectives)**
Intermediate Outcomes **(objectives)**
Longer-term Outcomes **(goal)**

Activities And Strategies
Accomplishments
Knowledge Attitudes Beliefs Skills Policies, Environment
Behaviors or Practices
Health or Disease

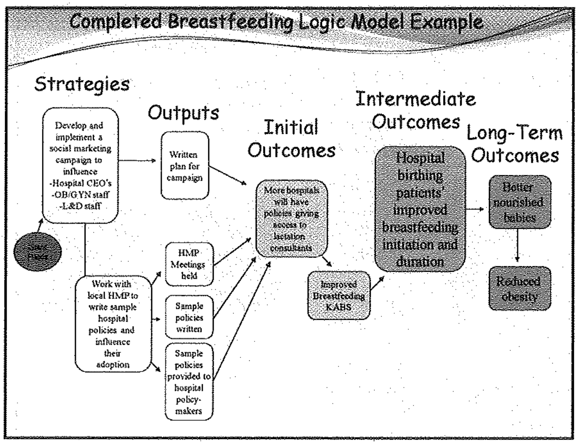
PE teachers trained using best practice guidelines
PE teachers completed training
Improved PE policies and practices
Students engage in more physical activity during PE classes
Improved middle school student fitness



Completed Adult Physical Activity Logic Model Example

Strategies
Outputs
Initial Outcomes
Intermediate Outcomes
Long-Term Outcomes

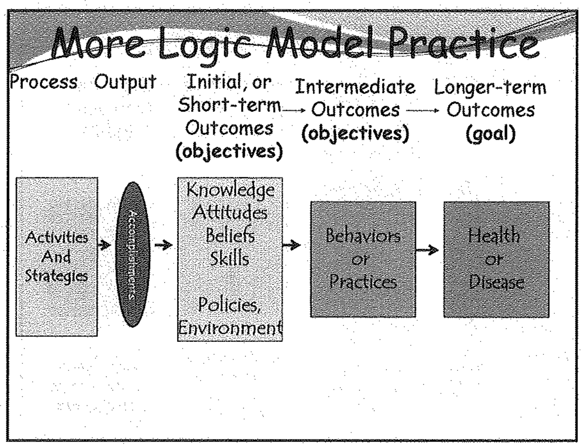
Recruit small business owners to steering group
Small business owners represented on steering group
Develop and conduct survey for small businesses to assess barriers and opportunities
Survey and survey results and recommendations
Develop one-year strategic plan
Written strategic plan
Complete strategic plan activities
Timely completion of strategic plan activities
Increase # of small businesses with flextime for PA policies in Portland
Increase PA in Portland small business workforce
Improved Fitness
Decreased CVD



Completed Breastfeeding Logic Model Example

Strategies
Outputs
Initial Outcomes
Intermediate Outcomes
Long-Term Outcomes

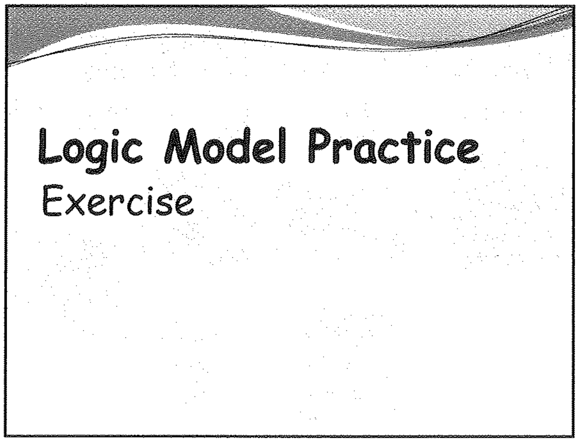
Start Here
Develop and implement a social marketing campaign to influence
-Hospital CEO's
-OB/GYN staff
-L&D staff
Written plan for campaign
Work with local HMP to write sample hospital policies and influence their adoption
HMP Meetings held
Sample policies written
Sample policies provided to hospital policy-makers
More hospitals will have policies giving access to lactation consultants
Improved Breastfeeding KABS
Hospital birthing patients' improved breastfeeding initiation and duration
Better nourished babies
Reduced obesity



More Logic Model Practice

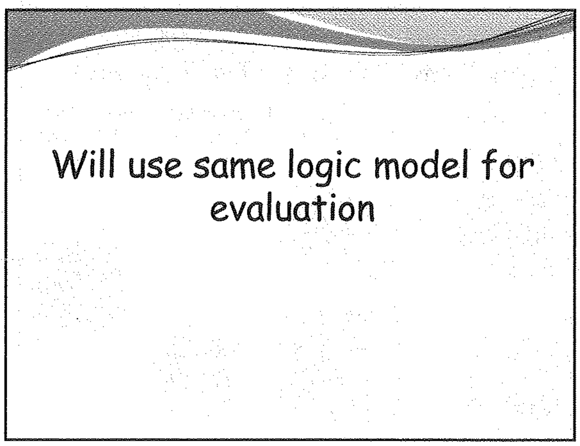
Process
Output
Initial, or Short-term Outcomes **(objectives)**
Intermediate Outcomes **(objectives)**
Longer-term Outcomes **(goal)**

Activities And Strategies
Accomplishments
Knowledge Attitudes Beliefs Skills Policies, Environment
Behaviors or Practices
Health or Disease



Logic Model Practice

Exercise



Will use same logic model for evaluation

Write Objectives

(SMART= Specific, Measurable, Attainable, Realistic, Time-bound)

- Who? (the population)
- What? (what specifically will you accomplish)
- How Much? (usually stated as a percentage)
- By When? (usually stated as a timeframe)

Objectives: Examples

- By March 15th 2011, conduct two field trips to local academic institutions for at least 50 at-risk adolescent girls in Middletown High School. (process)
- By June 1st 2011, 100% of all high schools in Cumberland County will comply with Maine's junk food and beverage marketing ban (initial)
- By May 2013 There will be a 10% decrease in cigarette smoking initiation among Middletown high school students as compared to May 2010. (intermediate)

Objectives

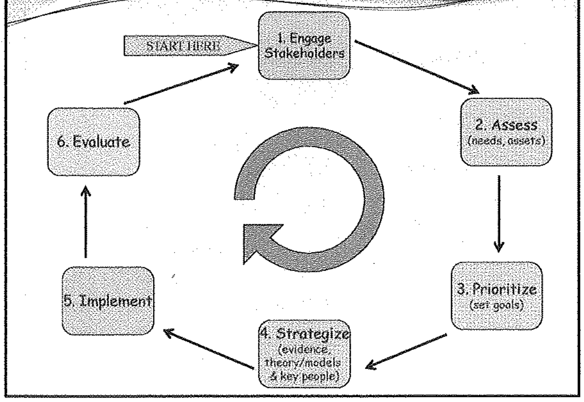
Exercise

-HIV prevention

-Youth smoking

Need measurable objectives for evaluation

Public Health Program Planning



START HERE

1. Engage Stakeholders

2. Assess (needs, assets)

3. Prioritize (set goals)

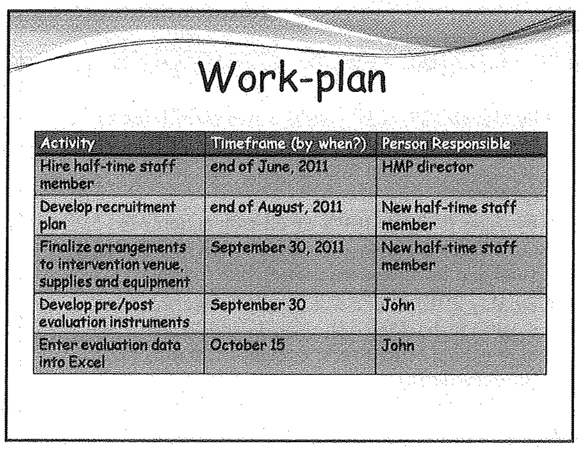
4. Strategize (evidence, theory/models & key people)

5. Implement

6. Evaluate

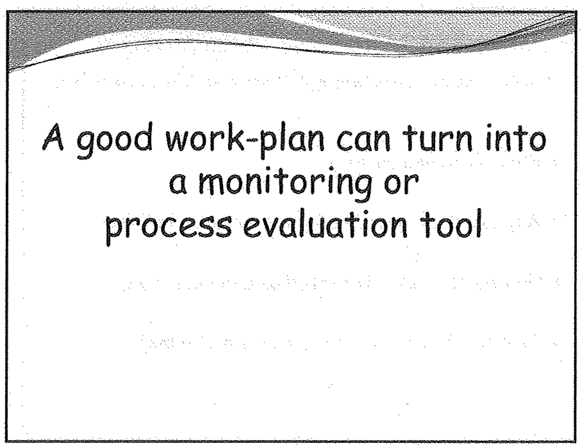
Implementation

- Implement at all ecological levels, if possible
- Specify timing and logistics (work-plan)
- Train staff
- Develop recruitment and retention plan (if appropriate)
- Pilot

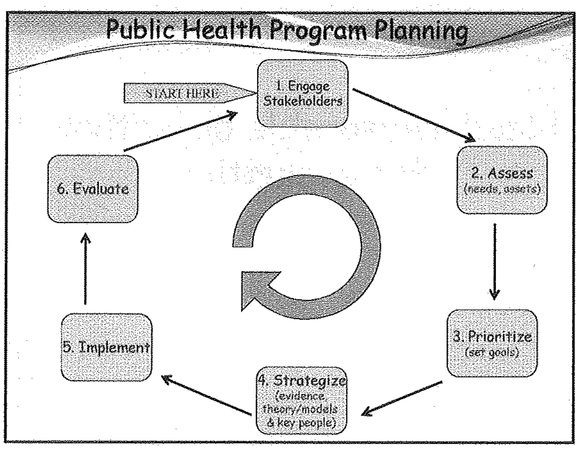


Work-plan

Activity	Timeframe (by when?)	Person Responsible
Hire half-time staff member	end of June, 2011	HMP director
Develop recruitment plan	end of August, 2011	New half-time staff member
Finalize arrangements to intervention venue, supplies and equipment	September 30, 2011	New half-time staff member
Develop pre/post evaluation instruments	September 30	John
Enter evaluation data into Excel	October 15	John

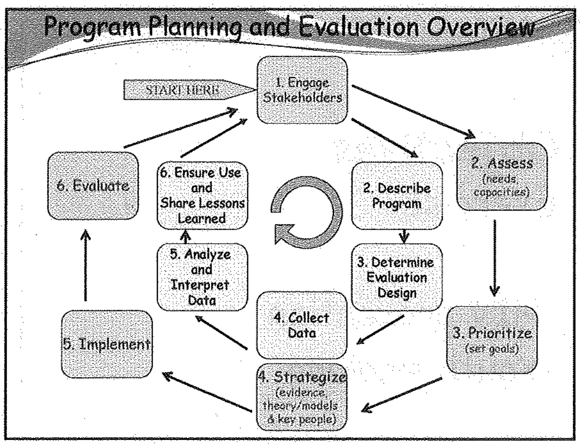


A good work-plan can turn into a monitoring or process evaluation tool



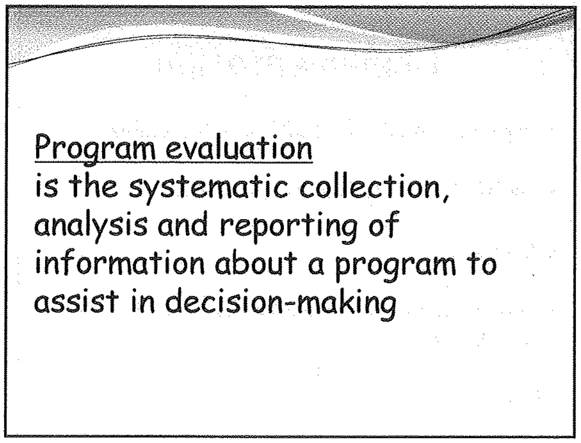
Public Health Program Planning

START HERE
1. Engage Stakeholders
2. Assess (needs, assets)
3. Prioritize (set goals)
4. Strategize (evidence, theory/models & key people)
5. Implement
6. Evaluate

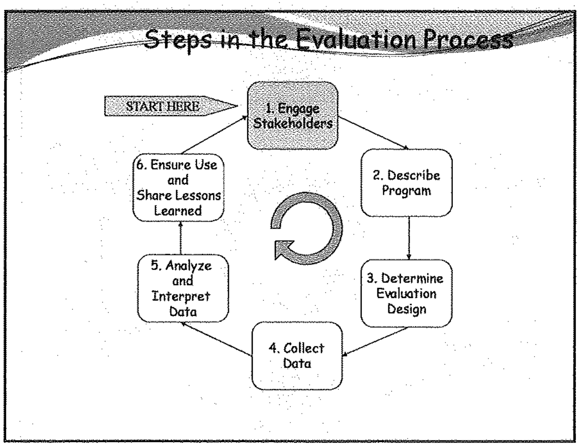


Program Planning and Evaluation Overview

START HERE
1. Engage Stakeholders
2. Assess (needs, capacities)
3. Prioritize (set goals)
4. Strategize (evidence, theory/models & key people)
5. Implement
6. Evaluate
2. Describe Program
3. Determine Evaluation Design
4. Collect Data
5. Analyze and Interpret Data
6. Ensure Use and Share Lessons Learned



Program evaluation is the systematic collection, analysis and reporting of information about a program to assist in decision-making



Steps in the Evaluation Process

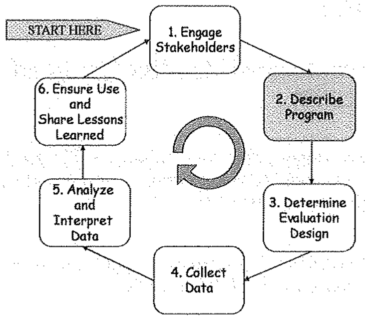
START HERE
1. Engage Stakeholders
2. Describe Program
3. Determine Evaluation Design
4. Collect Data
5. Analyze and Interpret Data
6. Ensure Use and Share Lessons Learned

Stakeholders?

People or organizations invested in the program, who have an interest in the results, and/or have a stake in what will be done with the results

- People involved in program operations
  (managers, staff, funders, coalition members)
- Those served or affected
  (patients, clients, advocacy groups, community members, elected officials)
- Users
  (policy makers, funders, taxpayers, general public, program critics)

Steps in the Evaluation Process



START HERE

1. Engage Stakeholders

2. Describe Program

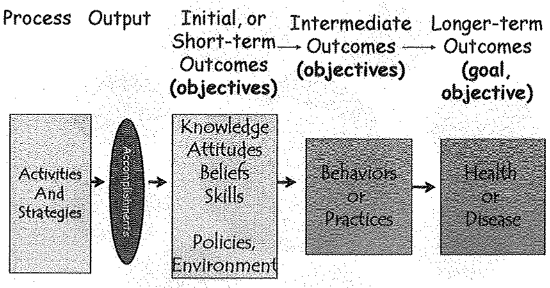
3. Determine Evaluation Design

4. Collect Data

5. Analyze and Interpret Data

6. Ensure Use and Share Lessons Learned

Logic Model



Process

Output

Initial, or Short-term Outcomes **(objectives)**

Intermediate Outcomes **(objectives)**

Longer-term Outcomes **(goal, objective)**

Activities And Strategies

Accomplishments

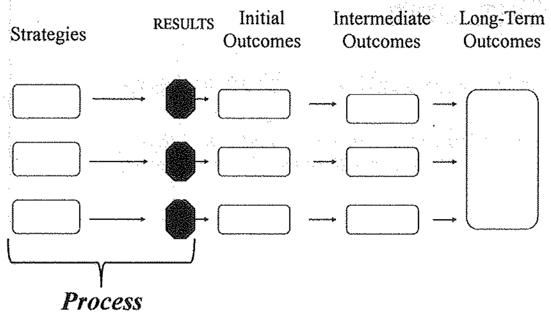
Knowledge Attitudes Beliefs Skills

Policies, Environment

Behaviors or Practices

Health or Disease

Program Overview Logic Model



Strategies

RESULTS

Initial Outcomes

Intermediate Outcomes

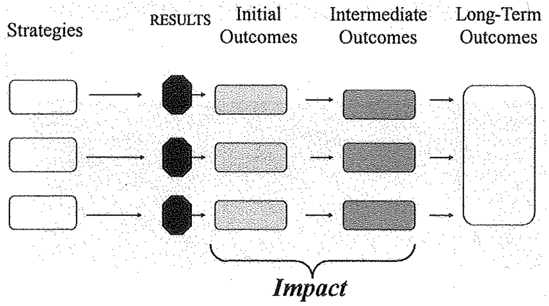
Long-Term Outcomes

Process

Process

- HOW a program is planned and/or implemented
- What activities were conducted?
  - What did you accomplish?
  - What services were actually provided?
- What materials did participants receive?
- What did people experience?

Program Overview Logic Model



Strategies

RESULTS

Initial Outcomes

Intermediate Outcomes

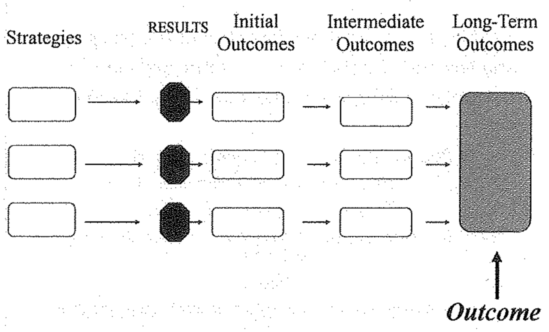
Long-Term Outcomes

Impact

Impact

- Initial or intermediate effects or benefits of a program
  - Did knowledge, attitudes, beliefs, skills (KABS), policies or environments change as a result of the program? **INITIAL**
  - Did behaviors or practices change as a result of the program? **INTERMEDIATE**

Program Overview Logic Model

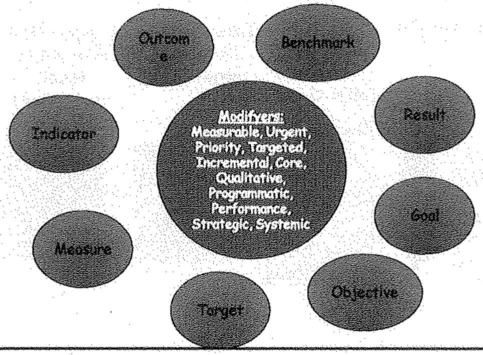


Outcome

- Longer term effects on health or disease
  - *what changes in injury, death, or disease (or cost) occurred (because of the program?)*

The Language Trap

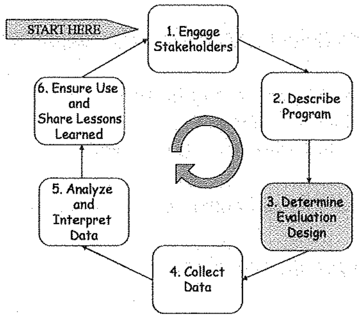
Too many terms. Too few definitions. Too little discipline


Outcom
e
Benchmark
Indicator
Modifyers:
Measurable, Urgent, Priority, Targeted, Incremental, Core, Qualitative, Programmatic, Performance, Strategic, Systemic
Result
Goal
Measure
Target
Objective

Choosing A Common Language

A condition of well-being for children, families or communities	A measure for which data are available, which helps quantify achievement	A measure of the effectiveness of program delivery
Outcome Result Goal Impact Other…	Indicator Benchmark Milestone Other…	Performance Measure Program Measure Process Measure Other…

Steps in the Evaluation Process


START HERE
1. Engage Stakeholders
2. Describe Program
3. Determine Evaluation Design
4. Collect Data
5. Analyze and Interpret Data
6. Ensure Use and Share Lessons Learned

First: Determine Evaluation Questions
(injury training example)

- **Process:**
  - What is the content of the training?
  - How are the trainees being selected?
  - What is the format of the training?
  - Are the sessions well attended?

Determine Evaluation Questions (training example)

- **Initial/intermediate outcomes:**
  - Has the client learned new skills?
  - Has the client used the new skills learned?

Determine Evaluation Questions (training example)

- **Longer-Term Outcomes:**
  - Are there less patient injuries because of the training?
  - Were *costs* reduced because of the training?

Determine Evaluation Design (to fit evaluation questions)

- Experimental
- Quasi-Experimental
- **Observational (time series, cross-sectional, case-studies)**

Quantitative Methods in Public Health Research

- Not all public health research lends itself to quantitative analysis.
- In cases where quantitative data have been collected from a reasonably sized sample, it is critical to choose the proper analytic methods.
- Poorly done analyses result in:
  - No publications
  - Publications that are criticized
  - Drawing the wrong conclusions

Types of Studies

- Experiments
  - Researchers manipulate the subjects' behavior in some way and monitor each group.
- Observational
  - Participants are simply observed by the researcher. The participants are not asked to behave, or do, anything differently.
- Which are you most likely to encounter in evaluation design?

Experiments

- In an experiment, we assign units to different groups with the goal of comparing them.
- ❖Experiments are particularly useful in that we can infer cause-and-effect relationships from them.
- To do this, we need to ensure that we have split the units into groups in a fair way.
- Example: What if you wanted to evaluate the effectiveness of an integrated physical/mental health intervention to a standard mental health only intervention?

Experiments

- How do you get subjects to participate in an experiment in the first place?
  - Offer them money.
  - Appeal to science.
- What's the problem with this?
- Those in your sample are likely to have participated, in part, because of their need for compensation if they're paid.
- If they're not paid, then a certain "type" of person may be more likely to participate.

Generalizability

- If your sample only consists of people with very strong views, then the decisions you make can only be applied to people with very strong views.
- Every type of person that you want to extend the results of your experiment to must be represented in a random way in your study.
- Once you have subjects recruited, how do you divide them into groups?

Randomization

- Randomizing group assignment is just as important as choosing a representative sample.
- If we have two (or more) groups, then we can randomly assign them as subjects to join the sample.
- This should balance out any underlying differences among subjects so that the groups are fair to compare.

Control Groups

- To see if one group differs from another, we need to compare two or more.
- Usually the group that we want to find an effect in is compared to a "standard" intervention that we call the control group.
- On the integrated services example, the mental health-only group would likely be the control.

Common Problems in Experiments

- Experimenter effects involve things such as the experimenter recording the results inaccurately, or interacting with the treatment group differently than the control group.
  - These are easily fixed with double blinding.
- Hawthorne effect: Common in medical studies where both researchers and patients more closely adhere to their treatment than in the "real world." This makes the treatment look more effective than it is in practice.

Observational Studies

- Why would we ever want to use an observational study over an experiment, especially since we cannot infer a cause-and-effect relationship effect from observational studies?
- This may not always be possible, or ethical.
- Why has it never been *proven* that smoking causes lung cancer?
- To prove cause-and-effect, we'd have to randomly assign people to smoke and not smoke.

Observational Studies

- There are three main types of observational studies.
  - Retrospective studies take a group and look backward in time to trace events.
  - Prospective studies follow a group over time and observe that group.
  - Case-control studies select the group to be studied based on the outcome in question.

Prospective or Retrospective?

- Retrospective studies rely on recall, or on records of past events, or secondary sources.
- For the same reason courts don't allow hearsay, using secondary sources is a risky business.
- Relying on recall can also be problematic if a lot of time has elapsed.
- Using records is better, although it depends on those keeping and reading them.
- Prospective studies are preferable if they can be done since they don't suffer from these pitfalls.

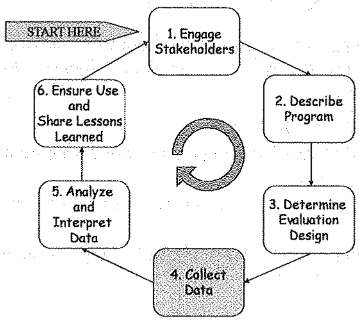
Case-control Studies

- The group of cases is found as a first step.
- Then a control group is selected that is as similar to the cases as possible in every way except for case status.
- Choosing the controls appropriately is very important
- Examples?

Case-control Studies

- Case-control studies are advantageous for a couple reasons:
  - If an outcome takes a long time to occur to is rare, then by selecting those that already have the outcome, we've eliminated the wait time and gotten a larger group of people.
  - By matching a group of controls to the cases based on important variables, discrepancies based on outside factors is reduced.

Steps in the Evaluation Process



START HERE

1. Engage Stakeholders

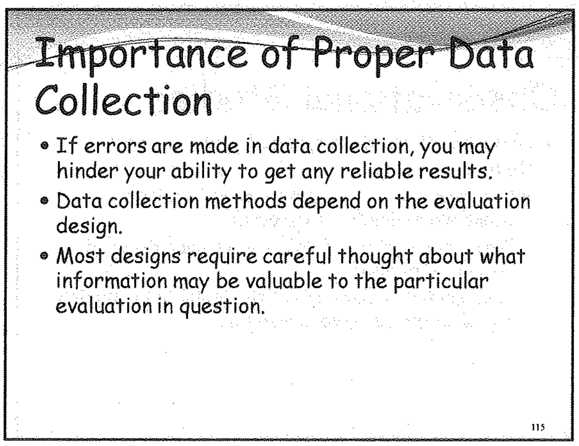
2. Describe Program

3. Determine Evaluation Design

4. Collect Data

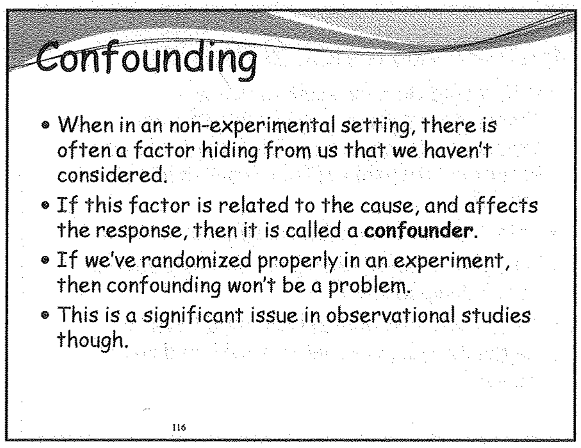
5. Analyze and Interpret Data

6. Ensure Use and Share Lessons Learned



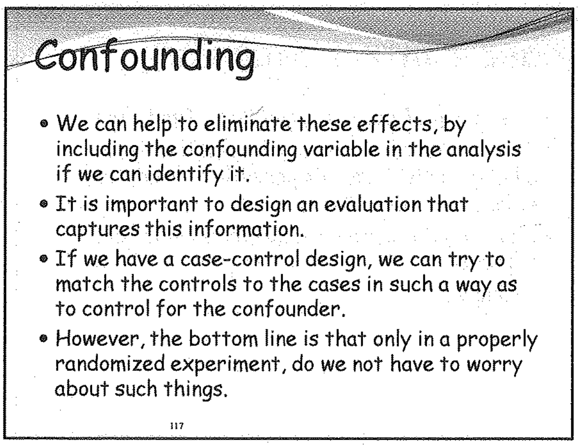
Importance of Proper Data Collection

- If errors are made in data collection, you may hinder your ability to get any reliable results.
- Data collection methods depend on the evaluation design.
- Most designs require careful thought about what information may be valuable to the particular evaluation in question.



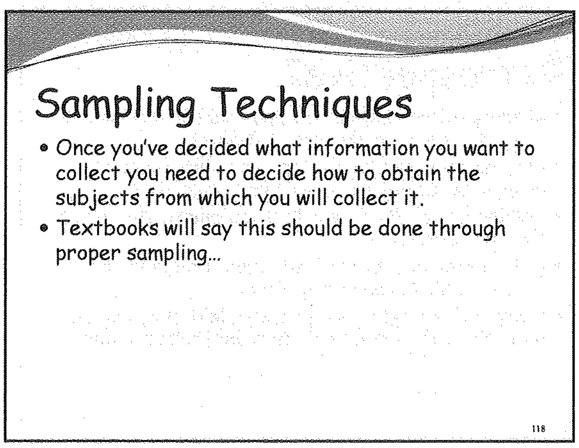
Confounding

- When in an non-experimental setting, there is often a factor hiding from us that we haven't considered.
- If this factor is related to the cause, and affects the response, then it is called a **confounder**.
- If we've randomized properly in an experiment, then confounding won't be a problem.
- This is a significant issue in observational studies though.



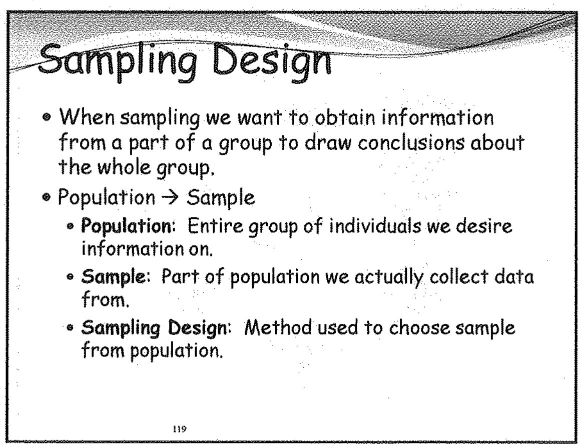
Confounding

- We can help to eliminate these effects, by including the confounding variable in the analysis if we can identify it.
- It is important to design an evaluation that captures this information.
- If we have a case-control design, we can try to match the controls to the cases in such a way as to control for the confounder.
- However, the bottom line is that only in a properly randomized experiment, do we not have to worry about such things.



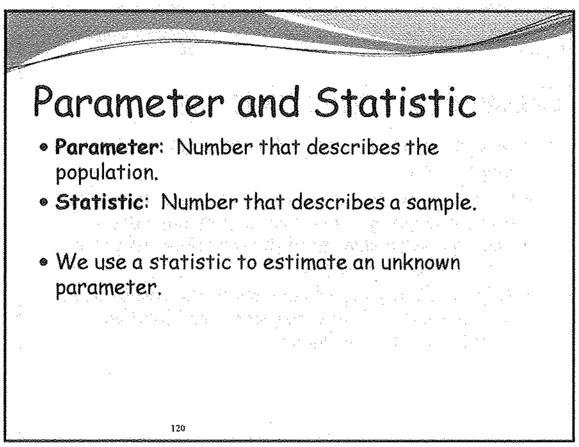
Sampling Techniques

- Once you've decided what information you want to collect you need to decide how to obtain the subjects from which you will collect it.
- Textbooks will say this should be done through proper sampling…



Sampling Design

- When sampling we want to obtain information from a part of a group to draw conclusions about the whole group.
- Population → Sample
  - **Population**: Entire group of individuals we desire information on.
  - **Sample**: Part of population we actually collect data from.
  - **Sampling Design**: Method used to choose sample from population.



Parameter and Statistic

- **Parameter**: Number that describes the population.
- **Statistic**: Number that describes a sample.

- We use a statistic to estimate an unknown parameter.

Simple Random Survey (SRS)

- In an SRS of size *n*:

1. Each individual has an *equal* chance of being chosen.
2. Every set of *n* individuals has an equal chance of being the sample chosen.

Stratified Samples

- **Basic Idea**: Sample important groups separately, then combine those samples.

1. Divide population into groups of similar individuals, called **strata**.
2. Choose a separate SRS within each strata.
3. Combine these SRS's to form the full sample.

Stratified Samples

- Strata for sampling are similar to blocks in experiments.
- Stratified sampling designs can provide more precise information than an SRS of the same size.
- For example, if all individuals within each stratum are identical, only need one individual from each stratum to perfectly describe the population.

Systematic Samples

- Choose a random starting point on a list.
- Add a fixed amount to that point.
- Repeat using same fixed amount.

Multistage Samples

- Basic Idea: Choose sample in stages.

- Often used for national surveys (U.S. households).
- Not practical to do SRS from list of all U.S. households (cost, inconvenience, time).

Multistage Samples

- To take a nationwide multistage sample:

1. Take sample from the 3000 counties in the U.S.
2. Take a sample of townships within each county chosen.
3. Take a sample of city blocks within each township chosen.
4. Take a sample of households within each city block.

- At each stage, take random sample (e.g., an SRS)

Sampling in Practice

- What is the most common type of sampling strategy seen in practice?
- The **convenience sample.**
- The biggest problem with the convenience sample...
- Generalizability

Surveys in Public Health

- Surveys are ubiquitous in many fields.
- Public health is no exception.
- Surveys can be powerful tools to gain information from an observational study.
- They are fraught with their set of difficulties in addition to those discussed in the observational study section.

Biases in Surveys

- **Selection Bias:** Some groups in population are over- or under-represented in sample.
- **Nonresponse Bias:** Nonrespondents may differ in important ways from respondents.
- **Response Bias:** e.g., wording of question, ordering of questions, telescoping in the recall of events.

1936 Literary Digest Poll

- Literary Digest had predicted the winner of every U.S. presidential election since 1916.
- In 1936, Literary Digest mailed questionnaires to 10 million people (25% of voters).
- 2.4 million people responded, the largest number of people ever replying to a poll.
- **Prediction:** Roosevelt 43%, Landon 57%
- **Actual Result:** Roosevelt 62%, Landon 38%

Selection and Nonresponse Bias

- **Selection Bias:** People surveyed came from telephone books, club memberships, mail order lists, automobile ownership lists.
- **Nonresponse Bias:** 76% did not respond.

- The Gallup Poll predicted Roosevelt's victory with a sample of 50,000 people.

Response Bias

- Wording of question can deliberately bias:
  - Do you favor, or do you not favor, increased restrictions on public smoking?
  - Do you favor Gestapo-like police tactics to prevent smoking in public?
  - Do you think smokers have the right to impose their filthy habits on the rest of us, polluting our precious air?

Response Bias

- Social Desirability:
  - Surveys of smoking underestimate the prevalence of smoking and do not match cigarette sales.
- Uninformed:
  - Survey by the American Jewish Committee on attitudes toward various ethnic groups.
  - "30% of respondents expressed an opinion about the Wisians…"

Survey Properties

- How do you develop a survey for program evaluation?
- Validity: Does it measure what you want it to measure?
- Reliability: Of you gave the survey a second time, would the results be consistent?
- We call these the "psychometric properties" of the survey.

Survey Properties

- In practice, it's best to use a survey that has already been developed and has good psychometric properties.
- Altering an existing survey is another option.
- Be aware that surveys for children and adolescents need to be validated for specific age ranges.
- If you develop your own the process of validating it should be handled by a professional.

Secondary Data

- Secondary data sources are commonly used in program evaluation.
- They can be a valuable tool for retrospective analysis.
- Be aware that data not collected for the purpose of evaluation may not always provide the best measure of the outcomes of interest.

Evaluation Study Designs

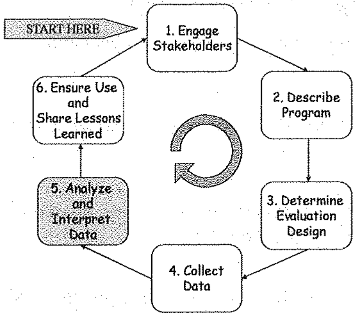
(where X=intervention; 0=data collection)

Study Design	Name
X 0	One shot case study
0 X 0	One group pretest/post test
X 0 0	Static group comparison group
0 0 X 0 0	Longitudinal Study/time series
R 0 X 0 R 0 0	Experiment with control group

Evaluation Data Collection Planning Chart

	Long-Term Outcome(s) (Health)	*Intermediate Outcome(s)* (Behavior)	*Initial Outcome(s)* (KABS, policies and environments)	Outputs	*Strategies* (What you're doing)
Measure(s) (Indicator)					
Data Source(s)					
Timing					

Steps in the Evaluation Process


START HERE

1. Engage Stakeholders

2. Describe Program

3. Determine Evaluation Design

4. Collect Data

5. Analyze and Interpret Data

6. Ensure Use and Share Lessons Learned

Data Analysis

- Proper data analysis is dependent on:
  - Type of evaluation done.
  - Method of data collection.
- Not all data lend themselves to quantitative analysis.
- Qualitative analysis is a useful complement to quantitative analysis or as a precursor to further data collection.

Common Problems in Observational Studies

- Extending results to a population that isn't represented by the sample.
- A third variable that isn't a confounder, but that interacts with the other variables. This can be fixed by only looking at subgroups with similar values of this variable.
- Ecological validity: By observing subjects in a non-standard setting, the subjects may alter their behavior for reasons unrelated to treatment.
- Causation…

"Smoking Causes Lung Cancer"

- SURGEON GENERAL'S WARNING: Smoking Causes Lung Cancer, Heart Disease, Emphysema, And May Complicate Pregnancy

---

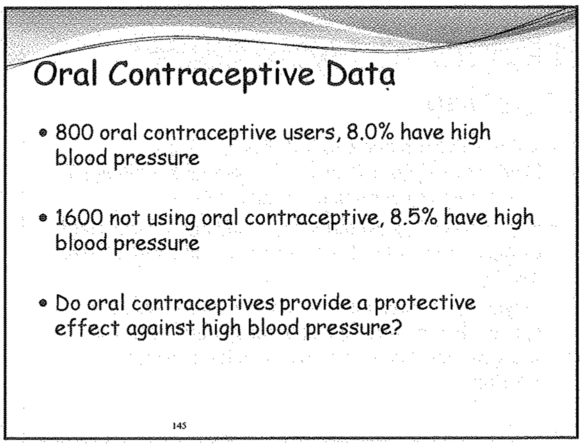
"Smoking Causes Lung Cancer"

- The association is strong.
- The association is consistent across many studies.
- High doses are associated with stronger responses.
- The alleged cause precedes the effect in time.
- The alleged cause is plausible.

Effect Modification

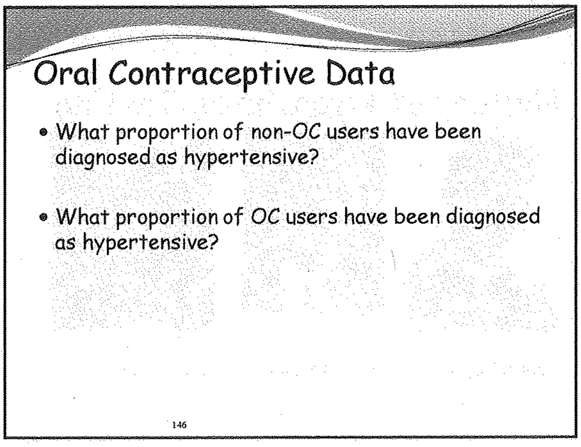
- A special type of confounding.
- Occurs when the effect of one variable on an outcome is altered by a second variable.

- **Simpson's Paradox**: Oral Contraceptive Data and Berkeley Graduate School Admissions Data.



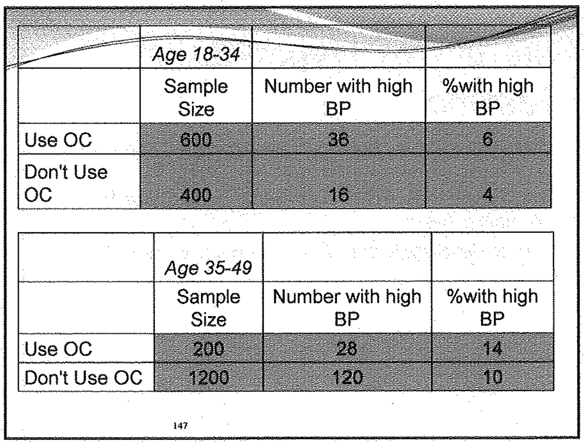
Oral Contraceptive Data

- 800 oral contraceptive users, 8.0% have high blood pressure
- 1600 not using oral contraceptive, 8.5% have high blood pressure
- Do oral contraceptives provide a protective effect against high blood pressure?



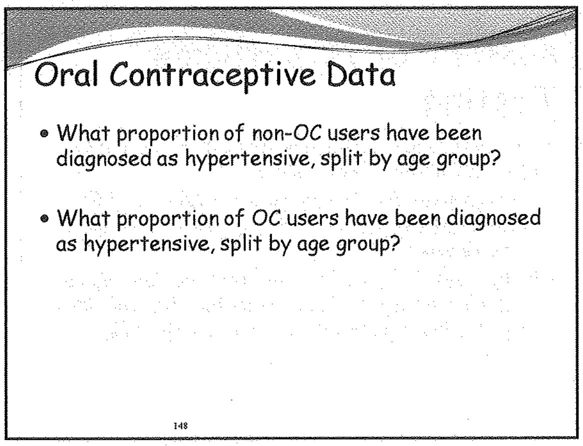
Oral Contraceptive Data

- What proportion of non-*OC* users have been diagnosed as hypertensive?
- What proportion of *OC* users have been diagnosed as hypertensive?



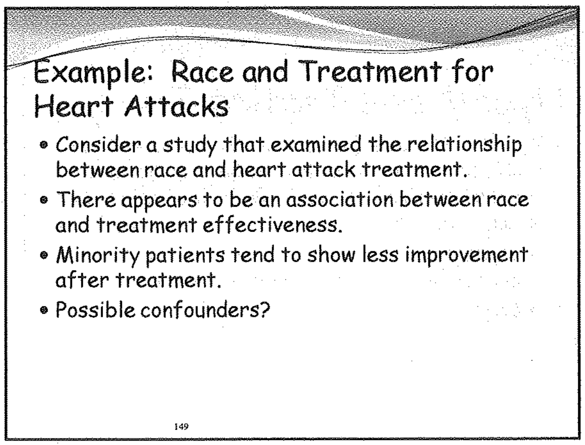
	Age 18-34		
	Sample Size	Number with high BP	%with high BP
Use OC	600	36	6
Don't Use OC	400	16	4

	Age 35-49		
	Sample Size	Number with high BP	%with high BP
Use OC	200	28	14
Don't Use OC	1200	120	10



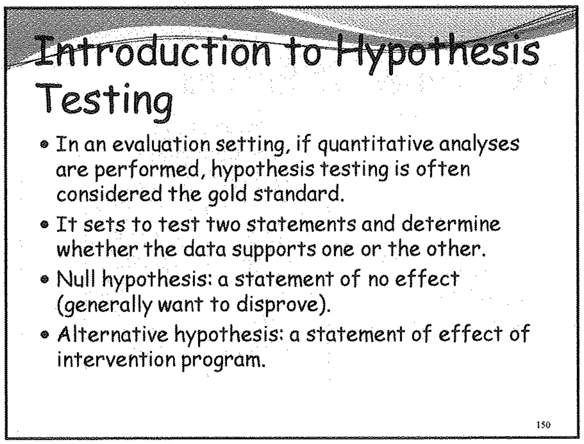
Oral Contraceptive Data

- What proportion of non-*OC* users have been diagnosed as hypertensive, split by age group?
- What proportion of *OC* users have been diagnosed as hypertensive, split by age group?



Example: Race and Treatment for Heart Attacks

- Consider a study that examined the relationship between race and heart attack treatment.
- There appears to be an association between race and treatment effectiveness.
- Minority patients tend to show less improvement after treatment.
- Possible confounders?

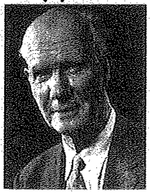


Introduction to Hypothesis Testing

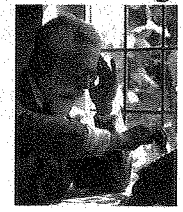
- In an evaluation setting, if quantitative analyses are performed, hypothesis testing is often considered the gold standard.
- It sets to test two statements and determine whether the data supports one or the other.
- Null hypothesis: a statement of no effect (generally want to disprove).
- Alternative hypothesis: a statement of effect of intervention program.

History of Hypothesis Testing

Jerzy Neyman



Egon Pearson



R.A. Fisher

History of Hypothesis Testing

- Roots can be traced to R.A. Fisher, Jerzy Neyman, and Egon Pearson (as well as Karl Pearson and William Gosset to a lesser extent).
- Fisher's work grew out of applications to the field of agriculture. He formulated the concept of the p-value.
- Neyman and Pearson took a more theoretical approach and together laid the groundwork for what we now know as statistical hypothesis testing.

History of Hypothesis Testing

- Pearson asked Neyman how he could be sure a set of data followed a bell curve..
  - Null hypothesis: The data are normally distributed
  - Alternative hypothesis: The data are not normally distributed
- Fisher had already addressed this problem saying that a large p-value simply indicated that the data provided inadequate information to reject H_0.

History of Hypothesis Testing

- Pearson and Neyman discussed the philosophical concepts of hypothesis testing carefully considering hypothesis formulation and how to compare tests.
- Little development has occurred since then.

Two Types of Errors

Truth about the population

		H_0 true	H_a true
Decision based on sample	Reject H_0	Type I error	Correct decision
	Accept H_0	Correct decision	Type II error

Hypothesis Testing Today

- W. Edwards Deming once said that students have such a hard time with the concept of hypothesis testing because "they might be trying to think."
- So who uses it?
  - The FDA
  - Most medical and public health researchers
  - Courts

Five Rules for Researchers

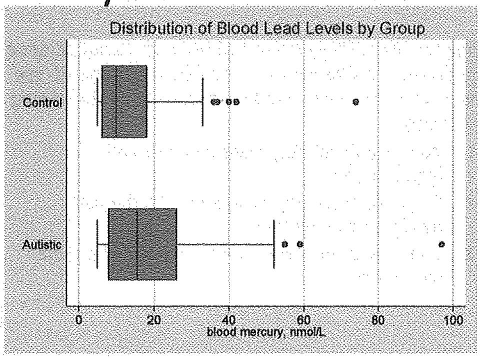
Motivation: Blood Lead Levels in Autism

- Ip, et al. (2004) paper in *Journal of Child Neurology.*
- "Student's t-test was used to compare the age, mean blood mercury level, and mean hair mercury level between the autistic and the normal group. Chi-square testing was used to test the sex ratio and social class between the autistic and the normal group. A significance level of $P < .05$ was used for all analyses. "

Motivation: Blood Lead Levels in Autism

- Ip et al. found that blood lead levels "did not significantly differ (p = NS)" between autistic and non-autistic groups.
- Two psychologists (DeSoto and Hitlan) noted that the p-value didn't make sense given the summary statistics.
- Full dataset was published along with a separate reanalyses by each of Ip and DeSoto.
- Ip said that their conclusions didn't change.

Reanalysis of Data



Reanalysis of Blood Lead Levels

- Desoto throws outliers away (Who needs them anyway? They just mess up your data!)
- **#1: Don't throw outliers away without a reason.**
- They take issue with the non-significance of the new t-test p-value by Ip (p = 0.056).
- **#2: Report all p-values and don't be fooled by the magical significance level of 0.05.**

Treatment of Outliers

- You can safely throw outliers away for any one of three reasons:
  - The data was recorded incorrectly.
  - The data came about from a sampling error.
  - You are willing to explicitly state that you know nothing about the part of the population that produced the outlying value and have results that have excluded the influence of that portion of the population.

What Happened to the Ozone?

- One of the most famous examples of inappropriately tossing out outliers comes from the story of Nimbus 7.
- In 1985 British Antarctic Survey recorded astonishingly low ozone levels.
- The Nimbus 7 satellite had been recording these since 1976 but never reported them.
- Its algorithm was programmed to throw outliers away!

Outlier Removal's Effect on the Test Statistic

- Desoto argues that since the larger outlier was in the autistic group, any bias that may result would be biased toward the null.
- Their claim is that the value of the test statistic will decrease since the distance between the mean will decrease.
- This is not necessarily true since the standard error is also affected inflating the test statistic.

P-values

- Gigerenzer (2004) reported 90% of social science professors surveyed endorsed at least one of six incorrect statements about the p-value (assuming $p < 0.05$).
  1. You have absolutely disproved the null hypothesis.
  2. You have found the probability of the null hypothesis being true.
  3. You have absolutely proved your experimental hypothesis.

P-values

4. You can deduce the probability of the experimental hypothesis being true.
5. You know, if you decide to reject the null hypothesis, the probability that you are making the wrong decision.
6. You have a reliable experimental finding in the sense that if, hypothetically, the experiment were repeated a great number of times, you would obtain a significant result on 99% of occasions.

- All six statements are incorrect and make the p-value look more informative than it really is.

P-values

- It's a conditional probability... and it's not pleasant to define.
- It is the probability that you would observe a test statistic as extreme or more extreme than the one you did observe *if the null hypothesis is true.*
- Where did the 0.05 value come from?

The Use of p= 0.05

- Fisher said in 1926 that, "...it is convenient to draw the line at about the level at which we can say 'Either there is something in the treatment, or a coincidence has occurred such as does not occur more than once in twenty trials...'"
- "Personally [I prefer] to set a low standard of significance at the 5% point, and ignore entirely all results which fail to reach that level."

Too Much Power?

- **#3: Given a large enough sample size, ANY difference can be deemed statistically significant.**
- No sample is perfect in practice.
- Biases from imperfect sampling get exacerbated as the samples gets larger.
- Causes even small biases to cause statistical significance.

Estimation and Hypothesis Testing: Two Peas in a Pod

- **#4: Always report an estimate of the effect size (along with a confidence interval) when reporting p-values.**
- The association between physical activity topics being taught in a required health class and TV time was significant, but the 95% CI: (1.03, 1.06).
- Confidence intervals and hypothesis testing provide complementary information and should both be used.

Estimation and Hypothesis Testing: Two Peas in a Pod

- Many papers over the past 60 years have criticized the use of hypothesis testing without providing effect size information.
- Thompson (1996) suggests three reforms for published research.
  - Replace the term "significant" with "statistically significant"
  - Require an effect size be reported.
  - Perform internal replication analyses using a resampling method.

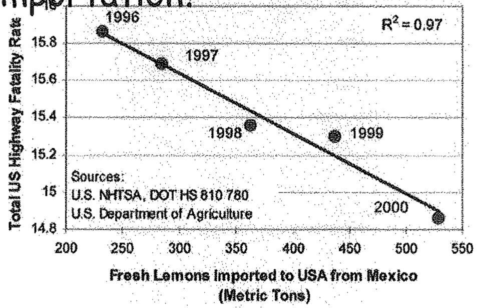
Association or Causation?

- **#5: Association (correlation) does not imply causation.**
- Unless a *well-designed* experiment has been conducted you cannot deduce causality from an association.
- The medical literature is rife with confounding results from observational studies.
- The effects of confounding can be complex.

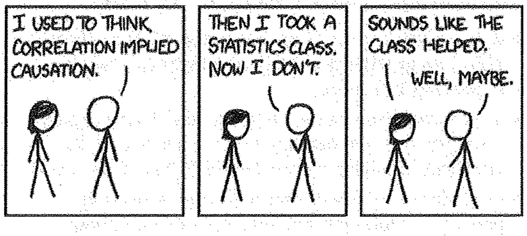
Association or Causation?

- Any time a variable is associated with the response and/or explanatory variable, confounding is present.
- There is no way to deduce causality in such a situation.

Should Auto Insurance Companies Subsidize Lemon Importation?



Sources:
U.S. NHTSA, DOT HS 810 780
U.S. Department of Agriculture


I USED TO THINK CORRELATION IMPLIED CAUSATION.
THEN I TOOK A STATISTICS CLASS. NOW I DON'T.
SOUNDS LIKE THE CLASS HELPED.
WELL, MAYBE.

Courtesy of xkcd.com

How to Turn Good Data into Bad Conclusions

Examples from the Literature

Intercessory Prayer

- Randomized controlled trial of remote intercessory prayer on the outcomes of patients admitted to the Coronary Care Unit (CCU).
- Published in 1999 in *Archives of Internal Medicine*.

Intercessory Prayer

- Patients were randomized to receive intercessory prayer (n=466) or no prayer (n=524) based on ID number.
- IRB exempted experiment from informed consent.
- Four outcome measures used:
  - Length of CCU stay
  - Length of hospital stay
  - MAHI-CCU raw score
  - MAHI-CCU weighted score

MAHI-CCU Score

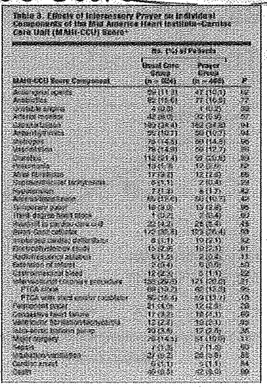


Table 3. Effects of Intercessory Prayer on Individual Components of the Mid America Heart Institute–Cardiac Care Unit (MAHI-CCU) Score*

Weighted MAHI-CCU Score

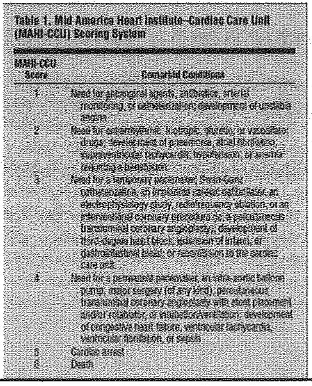


Table 1. Mid America Heart Institute–Cardiac Care Unit (MAHI-CCU) Scoring System

MAHI-CCU Score	Comorbid Conditions
1	Need for antianginal agents, antibiotics, arterial monitoring, or catheterization; development of unstable angina
2	Need for antiarrhythmic, inotropic, diuretic, or vasodilator drugs; development of pneumonia, atrial fibrillation, supraventricular tachycardia, hypotension, or anemia requiring a transfusion
3	Need for a temporary pacemaker, Swan-Ganz catheterization, an implanted cardiac defibrillator, an electrophysiology study, radiofrequency ablation, or an interventional coronary procedure (ie, a percutaneous transluminal coronary angioplasty); development of third-degree heart block, extension of infarct, or gastrointestinal bleed; or readmission to the cardiac care unit
4	Need for a permanent pacemaker, an intra-aortic balloon pump, major surgery (of any kind), percutaneous transluminal coronary angioplasty with stent placement and/or rotablator, or intubation/ventilation; development of congestive heart failure, ventricular tachycardia, ventricular fibrillation, or sepsis
5	Cardiac arrest
6	Death

Outcomes

- Out of 35 individual components that make up MAHI-CCU score, only one was found to be statistically significantly different (Swan-Ganz catheter, p = 0.03).
- Mean length of CCU stay difference (1.12 to 1.23, p = 0.28) in favor of prayer group.
- Mean length of hospital stay difference (5.97 to 6.48, p = 0.41) in favor of control group.
- Note: T-tests were used for all outcome comparisons.

Conclusions

- "Remote, intercessory prayer was associated with lower CCU course scores. This result suggests that prayer may be an effective adjunct to standard medical care."
- Raw MAHI-CCU score (3.0 to 2.7, p = .04).
- Weighted MAHI-CCU score (5.97 to 6.48, p = .04).

Do You Believe It?

- T-tests used on count data cannot be trusted.
- MAHI-CCU score is not validated.
- Weighted MAHI-CCU score is not validated.
- Multiple comparisons adjustment needed?
- What about differences in "background" prayer?
- Randomization not random.
- Ethical issues?

ESP Experiment

- The existence of psi has been of interest to experimental psychologists for decades.
- Some prior meta-analyses have shown statistically significant results for this existence (e.g., p = 0.0000003).
- In a soon-to-be published paper accepted by the *Journal of Personality and Social Psychology*, Daryl Bem discusses nine experiments conducted to test the existence of psi.

ESP Experiment

- Bem conducted a total of nine experiments:
  - Approach/avoidance
    - Precognitive Detection of Erotic Stimuli
    - Precognitive Avoidance of Negative Stimuli
  - Affective priming
    - Retroactive Priming I
    - Retroactive Priming II
  - Habituation
    - Retroactive Habituation I
    - Retroactive Habituation I
    - Retroactive Induction of Boredom
  - Facilitation of recall
    - Retroactive Facilitation of Recall I
    - Retroactive Facilitation of Recall II

ESP Experiment

- In 8 of these 9 experiments, Bem finds a significant p-value indicating ESP.
- For example, in experiment 1, he found that 53% of people correctly identified the future placement of an erotic picture.
- The conclusion is drawn that people "use psi information implicitly and nonconsciously to enhance their performance in a wide variety of everyday tasks."
- Let's take a closer look.

ESP Experiment

- Are Bem's experiments explanatory or confirmatory?
  - In experiment 1, Bem not only tested erotic pictures but also neutral, negative, positive, and romantic but not erotic pictures.
  - In experiments 5 and 6, tests were separated by gender despite no prior hypothesis for doing so.
  - In experiment 3, the latency times were transformed using two different transformations despite no need to do so.

ESP Experiment

- Confusion of the transposed conditional
  - Just because P(observed data given there no ESP) is small, that does not mean that P(no ESP given the observed data) is small.
  - A similar confusion exists in diagnostic testing between a test's sensitivity and it's positive predictive value.
- Lack of understanding of the information contained (and not contained) in the p-value.
  - Establishing how likely the data are under different, specific, alternatives would help.

Guidelines for Confirmatory Research

(Wagenmakers et al, 2011)

- Fishing expeditions should be prevented by selecting participants and items before the confirmatory study takes place.
- Transformations should only be applied if decided on beforehand.
- Analytic strategies should be determined before the experiment begins.
- It may be useful to consider multiple hypotheses when calculating p-values.

Antidepressants

- In 1998, Kirsch and Saperstein published an article in *Prevention & Treatment* declaring that antidepressants were no more effective than placebo.
- Used 19 FDA-approved experiments via a meta-analysis (n=2318).
- Found that placebo effectiveness was 75% that of active drugs.

More on Antidepressants

- This finding was replicated in more recent research published by Fournier et al in a 2010 *JAMA* article.
- They also used a meta-analysis on 6 FDA-approved placebo-controlled trials (n=718).
- However they considered baseline depressive symptoms as a possible effect modifier.

More on Antidepressants

- Effect modifier is the epidemiologic term for interaction.
- It indicates that the effect of one factor is dependent on the level of a second factor.
- This was found to be the case in antidepressant trials, with those with more severe depression seeing more improvement from drugs than those with less severe depression.

Fruits, Vegetables, and Cancer

- Diet, lifestyle, and cancer data obtained from the European Prospective Investigation into Cancer and Nutrition cohort study.
- Diet data collected primarily via 24-hour recall form 521,448 men and women.
- Estimated cancer risks adjusted for lifestyle variables including smoking status and alcohol consumption.

Fruits, Vegetables, and Cancer

- Analyses were statistically sound.
- Conclusions
  - There was a small, and significant, inverse relationship between fruit and vegetable consumption and cancer incidence (HR: 0.97, 95% CI: (0.96, 0.99)).
  - Among women only, a small reduction in cancer risk was associated with high vegetable intake (HR: 0.98, 95% CI: (0.97, 0.99).

Fruits, Vegetables, and Cancer

- Conclusion: "A very small inverse association between intake of total fruits and vegetables and cancer risk was observed in this study. Given the small magnitude of the observed associations, caution should be applied in their interpretation."

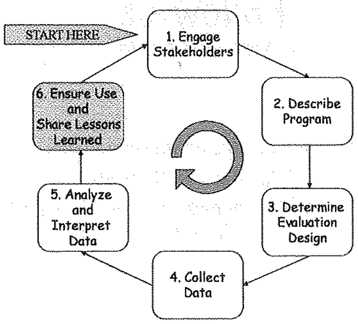
Potential Problems

- Recall bias!
- Small CIs due solely to huge sample size.
- Complex confounding issues.

Ways to Get the Most Out of Your Data

1. Use only validated measurement scales.
2. When performing exploratory analyses, be conservative when it comes to public reporting.
3. Don't transform data for no reason.
4. Don't throw away outliers.
5. Check all assumptions.
6. Report all potential limitations and sources of bias.
7. Graph your data.
8. Consult a statistician!

Steps in the Evaluation Process



START HERE

1. Engage Stakeholders

2. Describe Program

3. Determine Evaluation Design

4. Collect Data

5. Analyze and Interpret Data

6. Ensure Use and Share Lessons Learned

Five Elements to Ensure Use

- Recommendations
- Preparation
- Feedback
- Follow-up
- Dissemination

Make Recommendations

With stakeholders, come up with actions to consider as a result of the evaluation

Preparation

Outline steps to get ready for the use of the evaluation findings

Feedback

creates an atmosphere of trust

Communicate with everyone involved in the evaluation at all stages of the evaluation

- progress to date
- preliminary results
- opportunities to comment on evaluation decisions

Follow-up

proving support for users

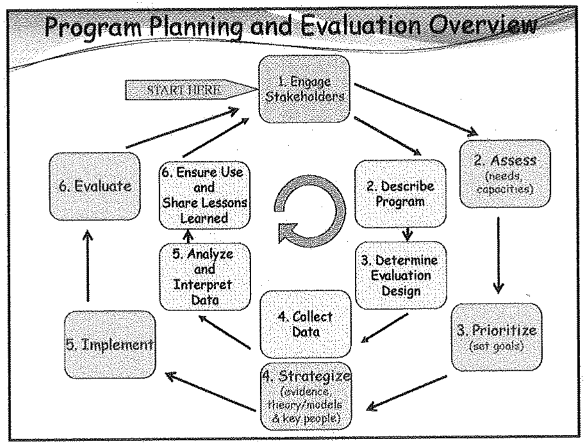
- Remind users of intended uses
- Help to prevent misuse
- Prevent lessons learned from becoming lost or ignored

Dissemination

(aiming for full disclosure)

The process of communicating evaluation procedures or lessons learned to relevant audiences in a timely, unbiased, and consistent manner

- Reports
- Mailings
- Websites
- Community forums
- Media
- Personal contacts
- List-serves
- Organizational newsletters



Program Planning and Evaluation Overview

START HERE

1. Engage Stakeholders

2. Assess (needs, capacities)

3. Prioritize (set goals)

4. Strategize (evidence, theory/models & key people)

5. Implement

6. Evaluate

2. Describe Program

3. Determine Evaluation Design

4. Collect Data

5. Analyze and Interpret Data

6. Ensure Use and Share Lessons Learned