

# Association of Pennsylvania Public Library Systems

Community Outcomes and Evaluation Workshop

August 21, 2015

Franklin & Marshall College

Center for

**OPINION RESEARCH**



# Overview

- Identifying Community Needs
- Traditional Evaluation Design
- Developmental Evaluation

# Identifying Community Need

- Defining Community
- Finding Partners Committed to Shared Measurement
- Finding Data
- Prioritizing Need

# Defining Community

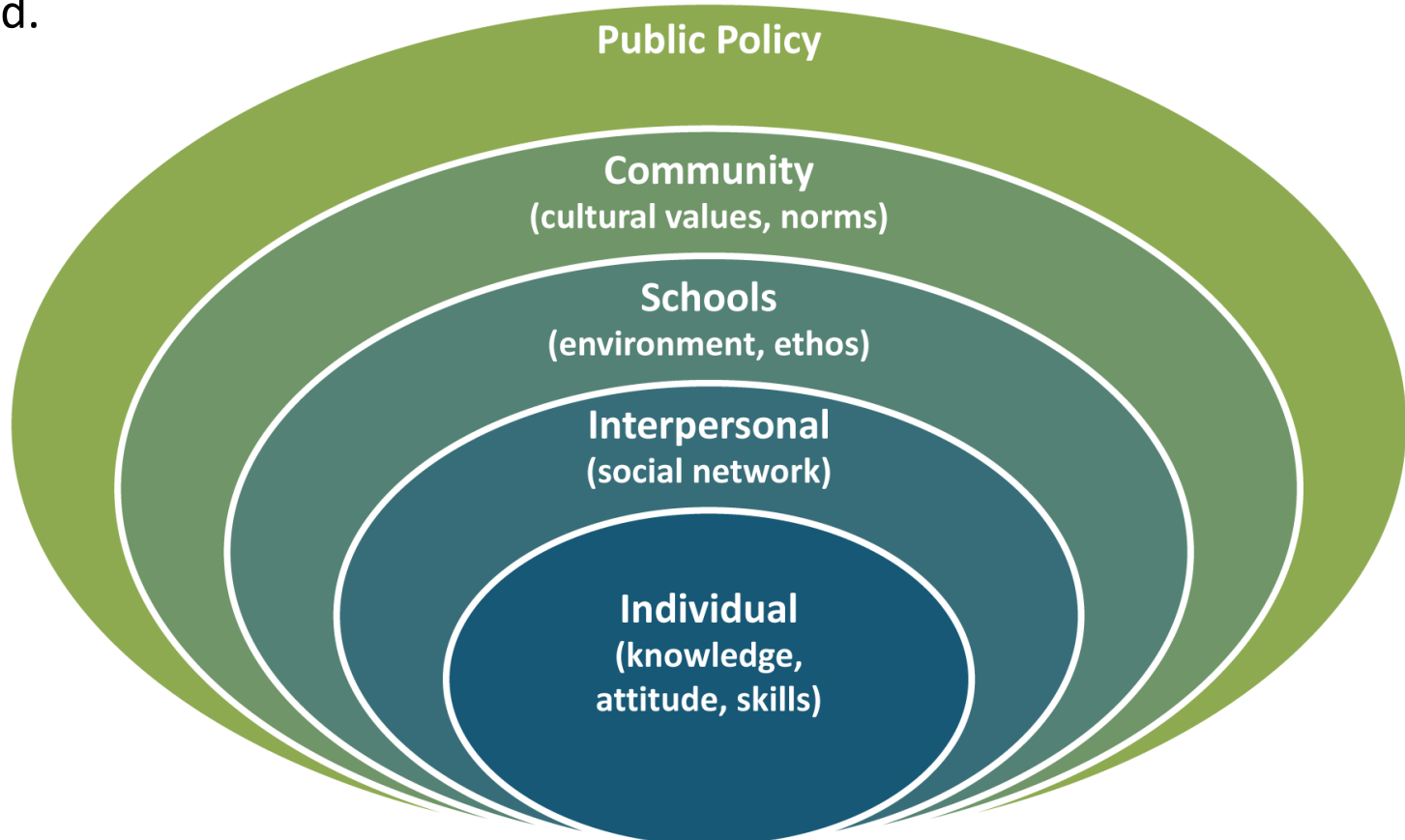
- Community - A group of individuals sharing one or more characteristics such as geographic location (e.g., a neighborhood), culture, age, or a particular risk factor. In the Guide to Community Preventive Services, for the purposes of evaluating whether interventions make communities healthier, we have chosen to apply the broadest possible use of “community.” <sup>1</sup>
- Community - a social group of any size whose members reside in a specific locality, share government, and often have a common cultural and historical heritage. <sup>2</sup>

• <sup>1</sup> CDC - The Community Guide - Glossary <http://www.thecommunityguide.org/about/glossary.html>

• <sup>2</sup> Dictionary Reference <http://dictionary.reference.com/browse/community>

# Shared Measurement

The implementation of complex community initiatives requires organizations to identify community partners who can work together to identify community need and solve community problems. The multiple causes underlying most complex problems require community partners from different segments of the community. The image below identifies the socioecological framework that shapes individual behavior and suggests areas where partners are needed.



# Group Exercise

## Defining Community

- Break into four groups (6 people per group)
  - Each group member should take on one of these roles (with a specific interest) during discussions:
    - Librarian (literacy)
    - Drug & Alcohol Agency (drug & alcohol issues)
    - United Way (poverty)
    - Community Action Program (poverty)
    - Catholic Charities (housing)
    - School District (graduation rates)
- Discuss and write down how you would define the community being targeted by this initiative
- Who makes up the community?
  - What are their key characteristics?

# Sources of Data on Community Need

- The Robert Wood Johnson Foundation (RWJ) County Health Rankings
  - <http://www.countyhealthrankings.org/app/pennsylvania/2015/overview>
- Community Commons
  - <http://www.communitycommons.org/maps-data/>
- The US Census Bureau – American Community Survey (ACS)
  - American Fact Finder
  - <http://factfinder.census.gov/faces/nav/jsf/pages/searchresults.xhtml?refresh=t>
- The Pennsylvania Department of Health - Epidemiologic Query and Mapping System (EpiQMS)
  - <https://apps.health.pa.gov/EpiQMS/asp/ChooseDataset.asp>
- The Pennsylvania Department of Education – Pennsylvania School Performance Profile
  - <http://paschoolperformance.org/SelectCounty>
- KIDS COUNT Data Center (A project of the Annie E. Casey Foundation)
  - <http://datacenter.kidscount.org/data#PA/2/0>
- Health Indicators Warehouse
  - <http://www.healthindicators.gov/>

# Prioritization

- ***What is prioritization?***

Prioritization is a process where an individual or group places a number of items in rank order based on their perceived or measured importance or significance.

Prioritization is generally a group process. Prioritizing issues is an important process, in that it assists an organization in identifying the issues on which it should focus its limited resources.

- ***Who is doing the prioritization?***

All participants usually have input into the prioritization process. Members of the prioritizing group need to be mindful that their own perceptions may be different from those around them. Often there is no clear right or wrong order to prioritizing, thus creating more difficulty in the prioritization process. That is especially true when trying to prioritize options that are unrelated or whose solutions are very different.

- ***Which method should be used?***

The following slides describe prioritization methods and the strengths and weaknesses of each. Some methods rely heavily on group participation, whereas other methods are less participatory and are more focused on baseline data. It is important to remember that no one method is best all of the time and each method can be adapted to suit the particular needs of a given community or group.



# Prioritization in Practice\*

Nominal Group. This method is useful in the early phases of prioritization when there exists a need to generate a lot of ideas in a short amount of time and when input from multiple individuals must be taken into consideration.

Multi-voting. Typically used when a long list of health problems or issues must be narrowed down to a top few

Strategy Grids. Strategy grids facilitate agencies in refocusing efforts by shifting emphasis towards addressing problems that will yield the greatest results.

The Hanlon Method. Though complex, the Hanlon Method is best when the desired outcome is an objective list of health priorities based on baseline data and numerical values.

Prioritization Matrix. A prioritization matrix is one of the more commonly used tools for prioritization and is ideal when health problems are considered against a large number of criteria or when an agency is restricted to focusing on only one priority health issue.

# Multi-voting Technique

Multi-voting is a quantitative tool used when a long list of health issues needs to be narrowed down. This can be accomplished in any manner where you can quickly tabulate votes such as hand-raising or using wireless voting technology. If you choose to use this technique, this sequence should be followed:

1. Round 1 Vote – Each participant votes for their highest priority items.
2. Update List – Health problems with highest votes remain on the list (problems with votes equivalent to or more than 50% of engaged participants).
3. Round 2 Vote – Each participant votes for their highest priority item from condensed list (votes per person limited to half the number of items remaining).
4. Repeat – Process repeated until list narrowed down to desired number of health priorities.

# Multi-voting Cont.

## Example of Three Round Multi-Voting Technique

Health Indicator	Round 1 Vote	Round 2 Vote	Round 3 Vote
Collect and maintain reliable, comparable, and valid data	√√√√	√√	
Evaluate public health processes, programs, and interventions.	√√√√√	√√√√	√√√√√
Maintain competent public health workforce	√√		
Implement quality improvement of public health processes, programs, and interventions	√√√√	√√	
Analyze public health data to identify health problems	√√		
Conduct timely investigations of health problems in coordination with other governmental agencies and key stakeholders	√√		
Develop and implement a strategic plan	√√√√√	√√√√	√√
Provide information on public health issues and functions through multiple methods to a variety of audiences	√√		
Identify and use evidence-based and promising practices	√√		
Conduct and monitor enforcement activities for which the agency has the authority	√		
Conduct a comprehensive planning process resulting in a community health improvement plan	√√√√√	√√√√	√√
Identify and implement strategies to improve access to healthcare services	√√√		√√

Red = Round 1 Elimination Green = Round 2 Elimination Blue = Round 3 Elimination

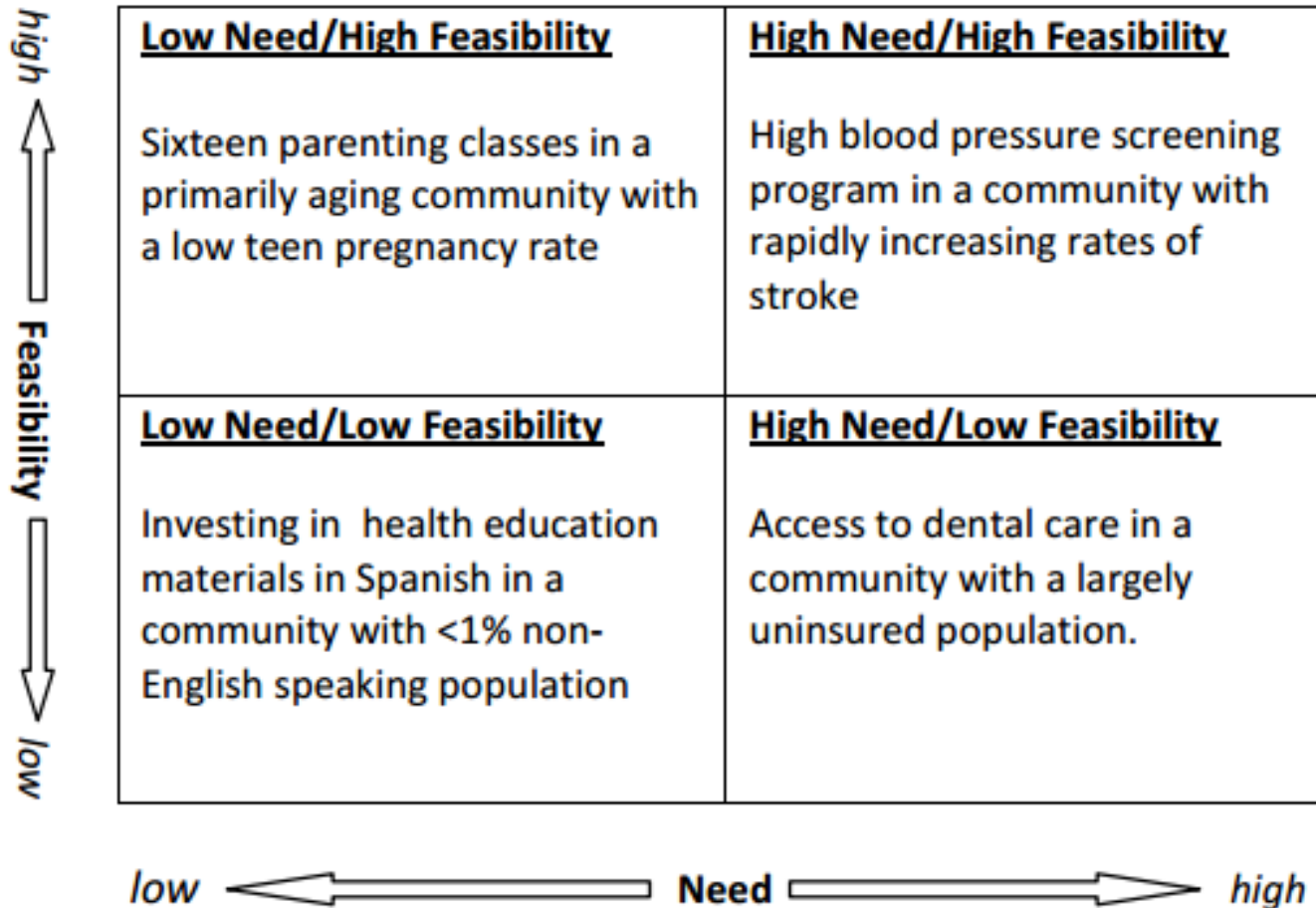
# Strategy Grids

## Step-by-Step Instructions:

1. **Select criteria** – Choose two broad criteria that are currently most relevant to the agency (e.g. 'importance/urgency,' 'cost/impact,' 'need/feasibility,' etc.). Competing activities, projects or programs will be evaluated against how well this set of criteria is met. The example strategy grid below uses 'Need' and 'Feasibility' as the criteria.
2. **Create a grid** – Set up a grid with four quadrants and assign one broad criteria to each axis. Create arrows on the axes to indicate 'high' or 'low,' as shown below.
3. **Label quadrants** – Based on the axes, label each quadrant as either 'High Need/High Feasibility,' 'High Need/Low Impact,' 'Low Need/High Feasibility,' 'Low Need/Low Feasibility.'
4. **Categorize & Prioritize** - Place competing activities, projects, or programs in the appropriate quadrant based on the quadrant labels. The example below depicts 'Need' and 'Feasibility' as the criteria and items have been prioritized as follows:
  - *High Need/High Feasibility* – With high demand and high return on investment, these are the highest priority items and should be given sufficient resources to maintain and continuously improve.
  - *Low Need/High Feasibility* – Often politically important and difficult to eliminate, these items may need to be re-designed to reduce investment while maintaining impact.
  - *High Need/Low Feasibility* – These are long term projects which have a great deal of potential but will require significant investment. Focusing on too many of these items can overwhelm an agency.
  - *Low Need/Low Feasibility* – With minimal return on investment, these are the lowest priority items and should be phased out allowing for resources to be reallocated to higher priority items.

# Strategy Grids, Cont.

**Strategy Grid**



# The Hanlon Method

This is a quantitative tool that objectively ranks specific health problems based on the criteria of seriousness, magnitude and effectiveness. Below is a brief description of how to use this method.

1. Give each health problem a numerical rating on a scale of 0-10 for each of the three criterion shown in the columns. Below is an example of how this can be established.

Rating	Size of Health Problem (% of population w/health problem)	Seriousness of Health Problem	Effectiveness of Interventions
9 or 10	>25% (STDs)	Very serious (e.g. HIV/AIDS)	80% - 100% effective (e.g. vaccination program)
7 or 8	10% - 24.9%	Relatively Serious	61% - 80% effective
5 or 6	1% - 9.9%	Serious	41% - 60% effective
3 or 4	.1% - .9%	Moderately Serious	21 - 40% effective
1 or 2	.01% - .09%	Relatively Not Serious	5% - 20% effective
0	< .01% (Meningococcal Meningitis)	Not Serious (teen acne)	<5% effective (access to care)
<b>Guiding considerations when ranking health problems against the 3 criteria</b>	Size of health problem should be based on baseline data collected from the individual community.	Does it require immediate attention? Is there public demand? What is the economic impact? What is the impact on quality of life? Is there a high hospitalization rate?	Determine upper and low measures for effectiveness and rate health problems relative to those limits.

# The Hanlon Method, Cont.

2. **Apply the 'PEARL' Test** – Once health problems have been rated for all criteria, use the 'PEARL' Test to screen out health problems based on the following feasibility factors:
  - Propriety – Is a program for the health problem suitable?
  - Economics – Does it make economic sense to address the problem? Are there economic consequences if a problem is not carried out?
  - Acceptability – Will a community accept the program? Is it wanted?
  - Resources – Is funding available or potentially available for a program?
  - Legality – Do current laws allow program activities to be implemented?
3. **Calculate priority scores** – Based on the three criteria rankings assigned to each health problem in Step 1 of the Hanlon Method, calculate the priority scores using the following formula:
$$D = [A + (2 \times B)] \times C$$
Where: D = Priority Score  
A = Size of health problem ranking  
B = Seriousness of health problem ranking  
C = Effectiveness of intervention ranking
4. **Rank the health problems**– Based on the priority scores calculated in Step 3 of the Hanlon Method, assign ranks to the health problems with the highest priority score receiving a rank of '1,' the next high priority score receiving a rank of '2,' and so on.

# McLean County Health Department - The Hanlon Method Example:

As a part of the Illinois Project for Local Assessment of Needs (IPLAN), a community health assessment and planning process, the McLean County Health Department (MCHD) used the Hanlon Method to prioritize health problems in the community. After determining the top eight health problems from the community health assessment data, MCHD used the Hanlon Method to establish the top three focus areas the agency should address. The following steps were taken to implement the prioritization process:

- 1. Rate against specified criteria – To rate each health problem, MCHD used the following considerations for each Hanlon criterion. Table 3.2 illustrates the top three of the eight health problems and corresponding ratings for each criterion.
  - Size of the problem – the percentage of the population with the problem, with an emphasis on the percentage of the population at risk for the problem
  - Seriousness of the problem – morbidity rates, mortality rates, economic loss, and the degree to which there is an urgency for intervention
  - Effectiveness of the intervention – the degree to which an intervention is available to address the health problem
- 2. Apply the ‘PEARL’ test – After long discussion, all eight health problems passed the ‘PEARL’ test as the interventions for each problem were judged to be proper, economical, acceptable, feasible based on available resources, and legal.
- 3. Calculate the priority scores – Priority scores were calculated by plugging in the ratings from Columns A through B into the formula in Column D. The calculations of the top three priority scores are illustrated in Table 3.2

**Table 4.2: MCHD Hanlon Priority Scoring**

Health Problem	A Size	B Seriousness	C Effectiveness of Intervention	D Priority Score (A + 2B)C	Rank
Cancer	8	10	6	168	3
Cerebrovascular Disease	7	9	7	175	2
Heart Disease	10	10	7	210	1



# Prioritization Matrix

## Step-by-Step Instructions:

The following steps outline the procedure for applying a prioritization matrix to prioritize health issues. While working through each step, refer to **Table 4.1** below for a visual representation:

**Table 5.1: Example Prioritization Matrix**

	<b>Criterion 1 (Rating X Weight)</b>	<b>Criterion 2 (Rating X Weight)</b>	<b>Criterion 3 (Rating X Weight)</b>	<b>Priority Score</b>
<b>Health Problem A</b>	$2 \times 0.5 = 1$	$1 \times .25 = .25$	$3 \times .25 = .75$	2
<b>Health Problem B</b>	$3 \times 0.5 = 1.5$	$2 \times .25 = 0.5$	$2 \times .25 = 0.5$	2.5
<b>Health Problem C</b>	$1 \times 0.5 = 0.5$	$1 \times .25 = .25$	$1 \times .25 = .25$	1

- 1. Create a matrix** – List all health issues vertically down the y-axis (vertical axis) of the matrix and all the criteria horizontally across the x-axis of the matrix so that each row is represented by a health issue and each column is represented by a criterion. Include an additional column for the priority score.
- 2. Rate against specified criteria** – Fill in cells of the matrix by rating each health issue against each criterion which should have been established by the team prior to beginning this process. An example of a rating scale can include the following:  
  

**3 = criterion met well**  
**2 = criterion met**  
**1 = criterion not met**
- 3. Weight the criteria** – If each criterion has a differing level of importance, account for the variations by assigning weights to each criterion. For example, if ‘Criterion 1’ is twice as important as ‘Criterion 2’ and ‘Criterion 3,’ the weight of ‘Criterion 1’ could be .5 and the weight of ‘Criterion 2’ and ‘Criterion 3’ could be .25. Multiply the rating established in Step 2 with the weight of the criteria in each cell of the matrix. If the chosen criteria all have an equal level of importance, this step can be skipped.
- 4. Calculate priority scores** – Once the cells of the matrix have been filled, calculate the final priority score for each health problem by adding the scores across the row. Assign ranks to the health problems with the highest priority score receiving a rank of ‘1.’

# Lawrence-Douglas County Health Department:

## Example Prioritization Matrix

Prior to beginning the prioritization process, Lawrence-Douglas County Health Department (LDCHD) developed a decision-making team which was comprised of ten people including directors and coordinators from throughout the department. Next, upon completion of an agency self-assessment, LDCHD identified areas of weakness and created a list of three potential health indicators to improve upon, along with five criteria found to be most relevant in pinpointing which health indicator will prove to have the greatest impact on the needs of Lawrence-Douglas County. Once these variables were determined, the groundwork was in place and LDCHD was ready to use a prioritization matrix to weigh the identified health indicators against each criterion to make a final decision on a focus area for a QI project. The following steps were used to implement the process:

1. Create a matrix – LDCHD used the prioritization matrix shown in Table 4.2, with the chosen health indicators listed on the Y-axis and each criterion listed across the X-axis:

Proposed Area for Improvement Based on LHD Self-Assessment	Evaluative Criteria					Total Score
	Linkage to Strategic Vision (.25)	Do we need to improve this area? (.25)	What chance is there that changes we put into place will make a difference? (.5)	Likelihood of completion within the timeframe we have (.5)	Importance to Customer (customer is the one who would benefit; could be patient or community) (.75)	
Media strategy & Communications to raise public health awareness	3 X (.25)	4 X (.25)	4 X (.5)	3 X (.5)	3 X (.75)	7.5
Work within network of stakeholders to gather and share data and information	2 X (.25)	3 X (.25)	2 X (.5)	1 X (.5)	1 X (.75)	3.5
Continuously develop current information on health issues that affect the community	4 X (.25)	2 X (.25)	3 X (.5)	1 X (.5)	2 X (.75)	5

*\*Note: The numerical rankings in Table 3.1 are meant to serve as an example and do not reflect the actual rankings from LDCHD's prioritization process.*

# Lawrence-Douglas County Health Department: Example Prioritization Matrix Cont.

2. Rank each health indicator against criteria – Each member of the decision-making team was given this prioritization matrix and asked to fill it out individually based on the following rating scale:

4 = High priority

3 = Moderate priority

2 = Low priority

1 = Not priority

After completing the matrix, each team member individually discussed with the facilitators of the process the reasoning behind how the health indicators were rated.

3. Weight the criteria – Although LDCHD weighted each criterion equally, (i.e. each criterion was assigned a multiplier of 1) the numbers in red provide an arbitrary example of how an agency could assign weights to the criteria based on perceived importance. In this example, with multipliers of .5, 'Likelihood of making a difference' and 'Completion within timeframe' are weighted as twice as important as 'Linkage to strategic vision' and 'Need for improvement,' with multipliers of .25. With a multiplier of .75, 'Importance to customer' is weighted as three times as important.

4. Calculate priority scores – Final priority scores are calculated by adding the weighted scores across the row and recording it in the 'Total Score' column. Since LDCHD had the team complete multiple matrices, the total scores for each health indicator were added together to determine the final priority scores. With 'Media Strategies' receiving the highest priority score of 7.5, it was assigned a rank of '1' and identified as the highest priority health indicator.

**Conclusion** In a world with a growing number of health concerns, scarce resources, budget cuts, and conflicting opinions, it is very easy to lose sight of the ultimate goal - improving health outcomes. Often times these external forces drive the decision making process within a health department and make determining where to focus resources and time challenging. Prioritization techniques provide a structured approach to analyze health problems and solutions, relative to all criteria and considerations, and focus on those that will prove to have the greatest impact on the overall health of a community.

# Group Exercise

## Identifying Need

- Discuss and write down how you would identify need among the community targeted in the first exercise
  - Will you need to prioritize? What prioritization process would you use?
  - Visit the kids counts web site to identify data to support your priorities KIDS COUNT Data Center (A project of the Annie E. Casey Foundation)  
<http://datacenter.kidscount.org/data#PA/2/0>

# Evaluation Strategies

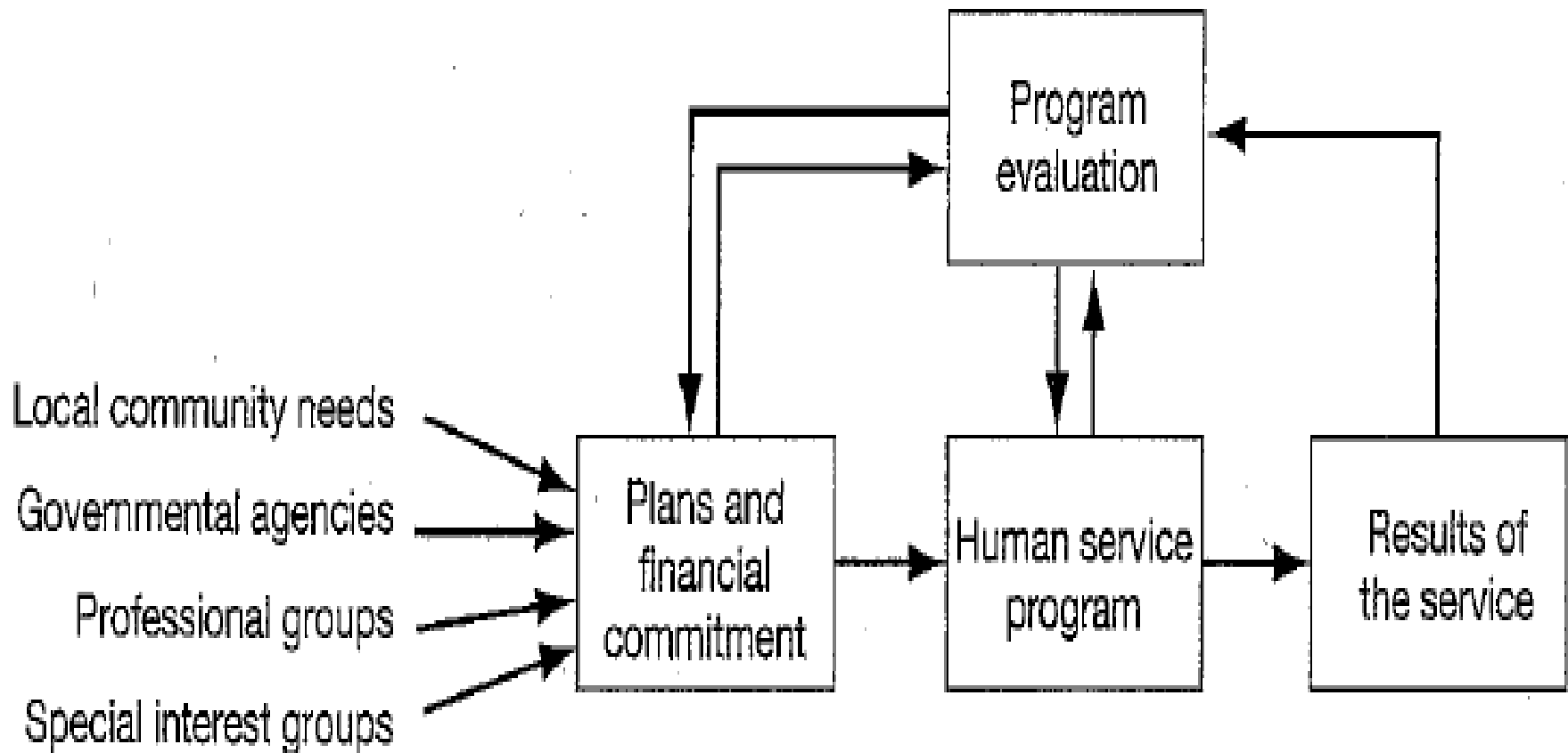
- Traditional Evaluation Practices
- Developmental Evaluation Practices

# Refresher on Traditional Evaluation

# Why evaluate?

- Many ask, why evaluate when there are scarce resources (time, money, expertise)? *Shouldn't I use these resources for providing services?*
- Evaluations can...
  - help determine whether services are deployed well and whether they do what they intend to do
  - Identify gaps in services
  - Help improve programs
  - Let funding agencies (and the broader community) know that their money is being well spent

# The primary purpose of the evaluation: feedback for improving the program





# Types of Evaluations

- Evaluation of need
  - Identify and measure unmet needs within an organization or community.
- Evaluation of process
  - Process evaluations document how a program has been delivered, who is being served, and whether the program is operating as expected.
- Evaluation of Impact
  - Impact evaluations show the effect a program has had on its intended targets.

# Process or Impact Evaluation?

- Program Goal: Provide and maintain safe environments in and around schools in SDOL
  1. Number of cameras placed in “high incident” areas
  2. Decrease in incidences of student violence and weapon possession by five percent per year
  3. Fidelity of anger management and conflict resolution group implementation
  4. Number of teachers, administrative staff and community partners that participate in safe school environment trainings
  5. Increase the percentage of teachers that feel safe in school by five percent per year

# Creating a Logic Model and Evaluation Plan

- A good evaluation plan begins with a good set of program goals and/or expected outcomes.
- Program goals can be translated into evaluation questions and measurable outcomes.

# Example of Logic Model

Needs and Gaps	Goals	SS/HS Element(s)	Objectives	Activities	Partners	Process Measures	Indicators and GPRA
<p>1. The PA Violence and Weapons Possession Report reported 182 incidents of violence or weapon possession including 41 assaults on students and 19 assaults on school staff.</p> <p>2. The SDOL reported 7,174 out of school suspensions assigned during the 2003-2004 school year.</p> <p>3. The Federal Bureau of Investigation reported 1,580 juvenile arrests in Lancaster city in 2002.</p> <p>4. The Federal Bureau of Investigation reported 6,435 crimes committed per 100,000 people in Lancaster city in 2003.</p>	To provide and maintain safe environments in and around schools in SDOL.	Element 1: Safe school environment.	<p>Short –Term:</p> <p>1. Develop a tracking system for School Resource Officer activities and incidences of school specific violence by September 2006.</p> <p>2. To train school-based community staff and school staff around school safety by September 2007.</p> <p>Long-Term:</p> <p>1. To increase the percentage of students who feel safe at schools, to and walking to and from school, at home and in the community.</p> <p>2. Provide staff development around creating a safe and positive school environment.</p>	<p>Short –Term:</p> <p>1. Implement School Resource Officers at 3 schools in SDOL.</p> <p>2. Implement Safe Passage program at 1 school in SDOL.</p> <p>3. Implement secure card access systems and security cameras at schools.</p> <p>Long-Term:</p> <p>1. Create a safe school environment supported by professional development focused on positive school culture.</p> <p>2. Create a safe school environment by creating a culture of community focused on safety and crisis prevention</p>	<p>Lancaster Police will hire SROs and will track data dealing directly with school and community violence.</p> <p>SDOL will provide training on school safety to 3 schools and track data related to school discipline. SDOL will hire Juvenile Justice Liaison to act as contact person for juvenile justice system.</p> <p>Probation and Parole will work with police and SDOL to transition students to and from juvenile justice system and school system.</p>	<p>1. Number and type of incidents handled by SROs at school sites.</p> <p>2. Number of schools with card access systems and cameras.</p> <p>3. Number of teachers, administrative staff and school-based staff that participate in safe school environment trainings.</p> <p>4. School policies implemented or rewritten to support safe school environments.</p> <p>5. Percentage of students who feel safe in school.</p> <p>(Safe Passage measures)</p>	<p>1. Performance Indicator for Element 1 and GPRA 1 for decreases in student violence: Decrease in incidences of student violence by 5% per year from baseline collected 2003-2004 <u>as measured by SDOL suspension rates, juvenile attest rates, PA Violence and Weapons possession report and Lancaster city crime statistics.</u></p>

# Example of Evaluation Plan

<i><b>Evaluation Questions</b></i>	<i><b>Evaluation Design</b></i>	<i><b>Methodology</b></i>	<i><b>Reporting</b></i>
<p><i>Was there a meaningful decrease (at least 14%) in the incidences of student violence and weapon possession district wide from baseline to June 2008?</i></p> <p><i>Does the presence of school resource officers deter student violence and weapon possession?</i></p>		<p><b>The School Violence and Weapons Possession Reporting System</b> will be used to monitor the incidences of student violence and weapon possession (see Appendix D).</p> <p>These data are reported annually by the district to the PA Department of Education in accordance with <b>Act 26 of 1995</b>, the Safe Schools Act. Data are submitted by June 30<sup>th</sup> of each year and are available at the district and school level.</p>	Data will be reported annually and will be analyzed at the district and school level.
<p><i>Was there a meaningful decrease (at least 5% per year) in the number of out-of-school suspensions over the course of the grant?</i></p>	<p>Interrupted Time Series with a Nonequivalent No-treatment Control Group Time Series</p> <p>O<sub>1</sub> O<sub>2</sub> O<sub>3</sub> X O<sub>4</sub> O<sub>5</sub> O<sub>6</sub> O<sub>1</sub> O<sub>2</sub> O<sub>3</sub> O<sub>4</sub> O<sub>5</sub> O<sub>6</sub></p>	<p>Out-of-school suspension data will be provided by the <b>District's Pentamation System</b>. Data will be collected at the end of each school year, Summers 2006, 2007, and 2008.</p>	Data will be reported annually and will be analyzed at the district and school level.
<p><i>How has student perception of school climate changed over the course of the grant?</i></p> <p><i>Do students feel safe in schools?</i></p> <p><i>Do students attending schools with school resource officers report feeling safer in school than students attending schools without school resource</i></p>		<p>All students, grades X-12, will be invited to participate in the <b>Annual SDoL Student Survey</b>. The Annual SDoL Student Survey will include measures of education quality, school climate, perception of school safety, students' social skills and emotional development (e.g., SSRS, SDQ), importance of higher education, etc. Data will be collected in Fall 2006,</p>	Survey results will be reported annually. Data will be analyzed at the district and school level.

# Where do you get data?

- Program Participants
- Providers (program staff)
- Independent Observers
- Existing records
- Any ideas?

# Developing Evaluation Questions

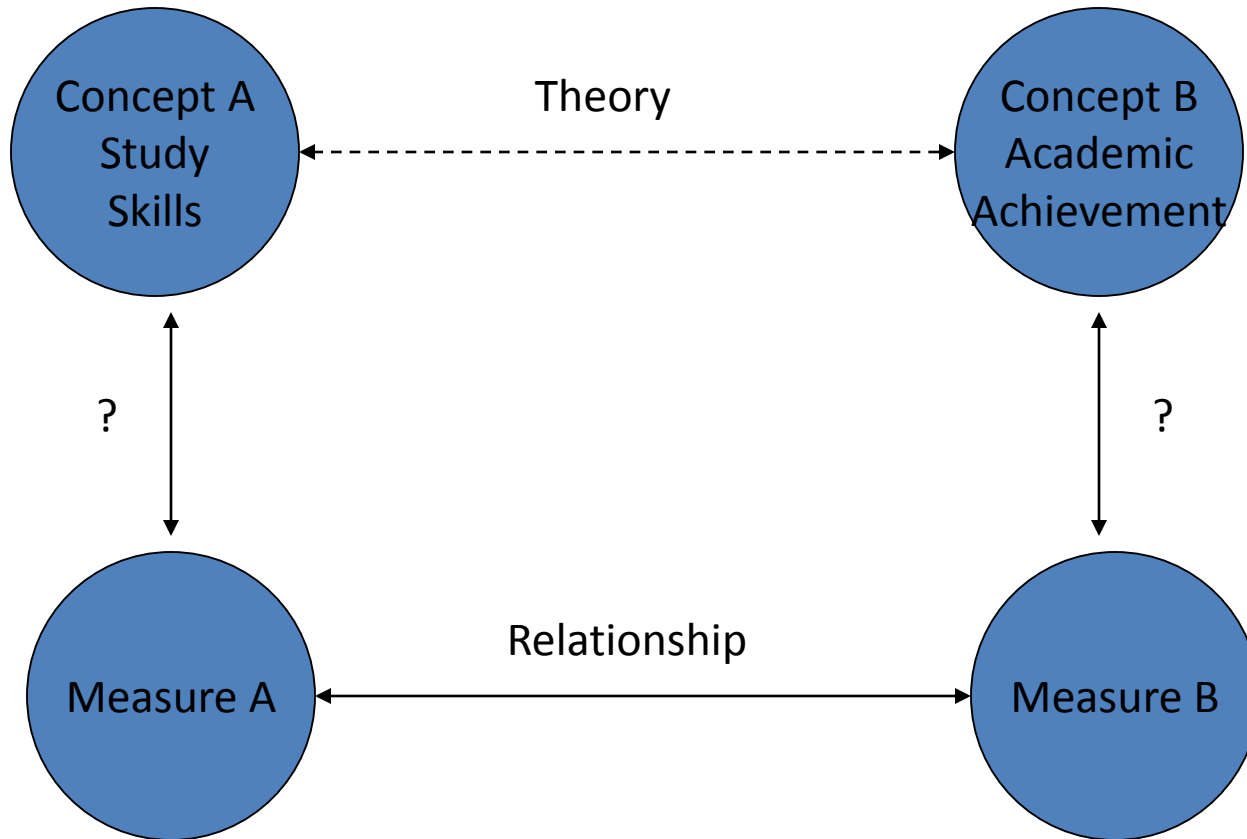
- Evaluation begins with a clear set of expectations about program goals.
- After determining goals you should answer two questions:
  1. What processes were (or should have been) instituted as a result of the program
  2. What expected behavioral changes/outcomes should result from the program.
- Answers to 1 and 2 will dictate your evaluation questions which will, in turn, suggest some ways of measuring each.

# Developing Evaluation Measures

- Short-term versus long-term
  - Too many program directors set themselves up to fail by focusing on goals that are too broad, long-term or have a small likelihood of being achieved.
- Example: Program Goal to improve student academic performance
  - Broad, long-term measure: Improve PA State Assessment Scores
  - Short-term measures: Improve study skills



# Developing Evaluation Measures



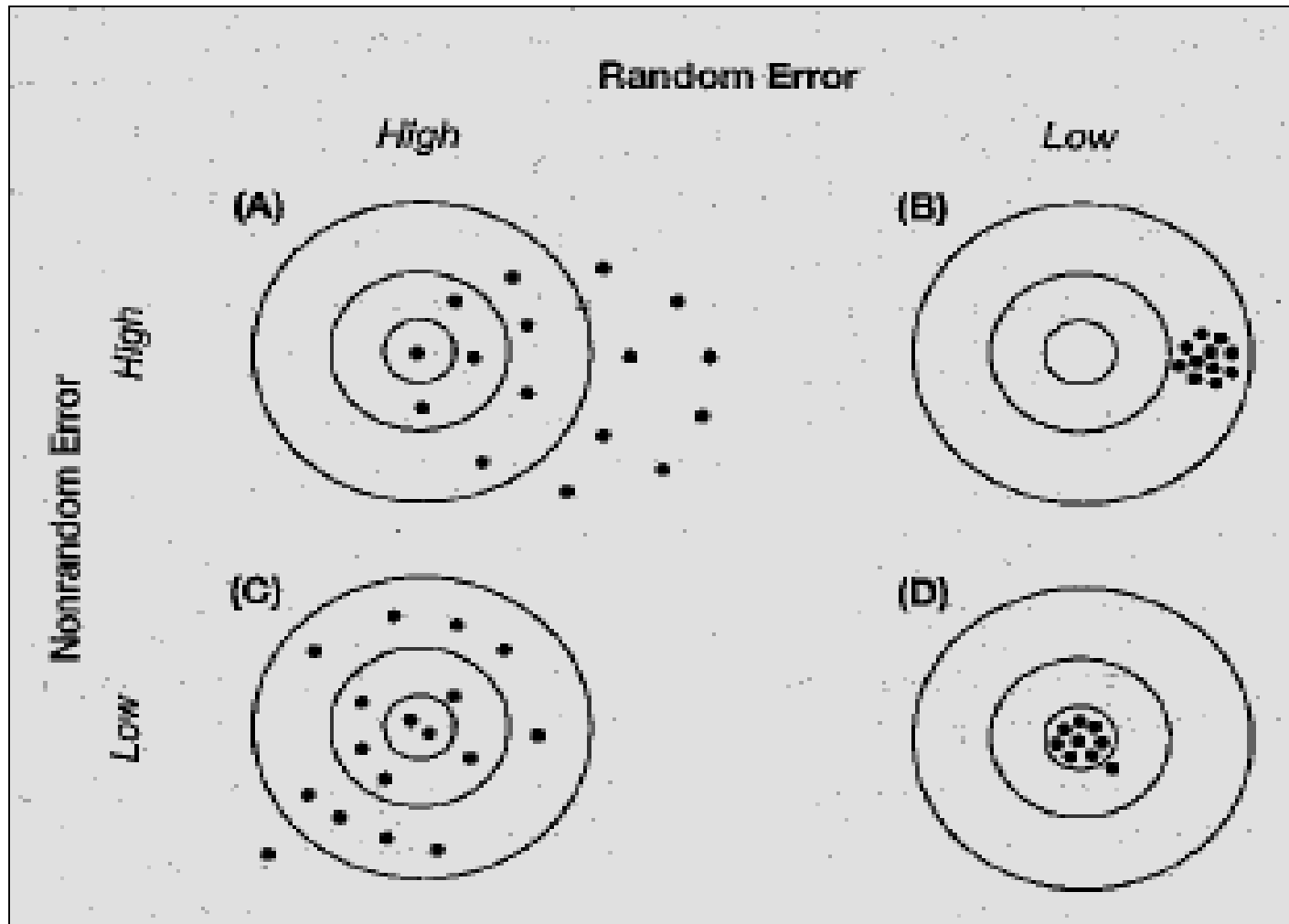
# Developing Evaluation Measures

- Sensitive
  - The measurements you make should contain as much information as possible about the attribute or behavior being measured, keeping in mind the goal of your evaluation.

# Developing Evaluation Measures

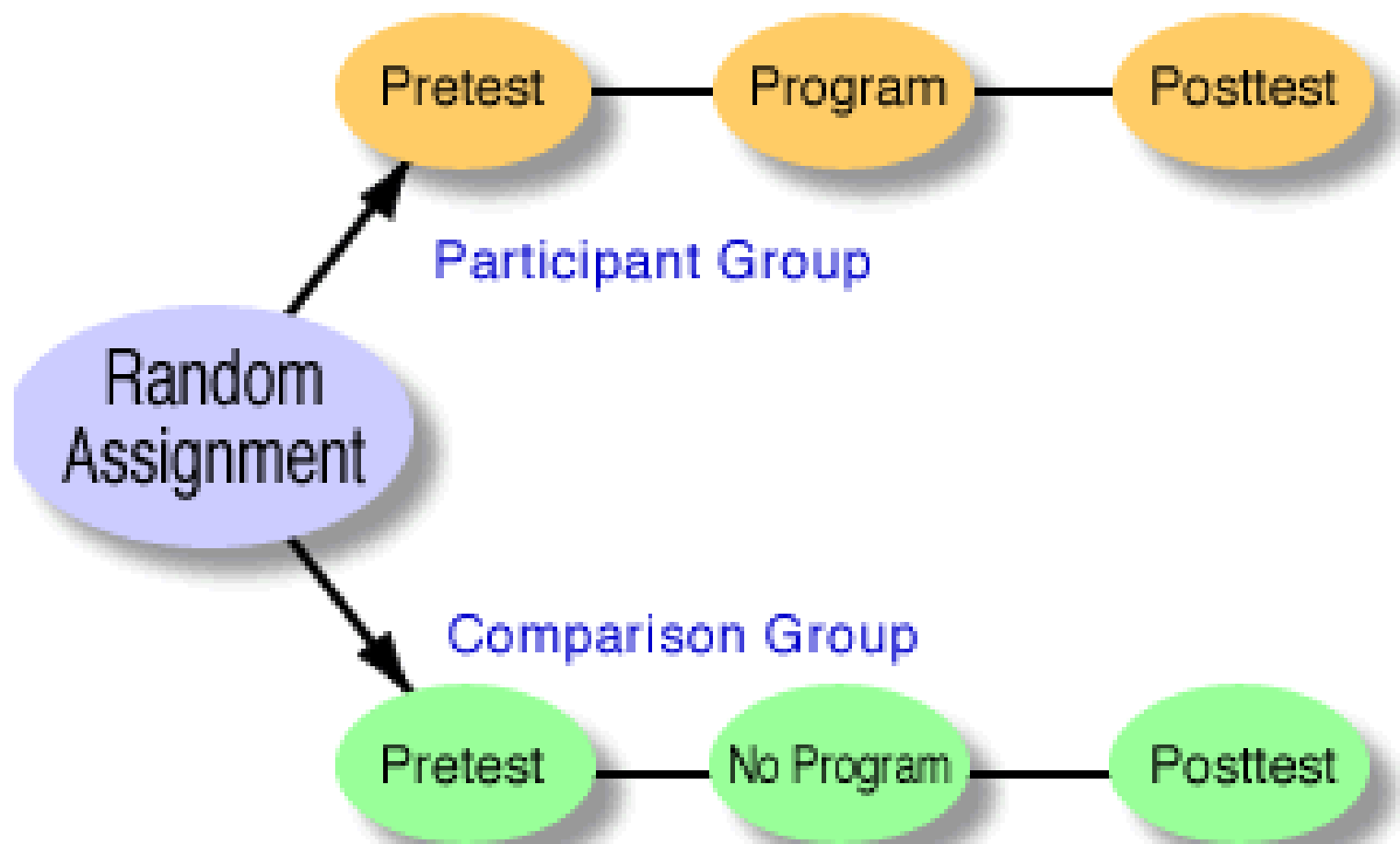
- Reliability
  - The extent to which an experiment, test, or any measuring procedure yields the same results on repeated trials
- Validity
  - The extent to which answers relate to some true value of what is being measured

# Reliability and Validity



# Choosing an Evaluation Design

- Experimental
  - Explores cause and effect relationships using random assignment to create equivalent groups
- Quasi-Experimental
  - Explores cause and effect relationships without using random, equivalent groups
- Non-experimental
  - Have little control over independent variables, units of analysis, or environment



# Group Exercise

## Logic Model

- Using the template provided, construct a logic model to identify the needs, goals, activities, and outcomes for one component of the Allegheny County initiative

# Intro to Developmental Evaluation



# Evaluation Goals and Expectations

1. Evaluations of complex, major initiatives are not experiments but part of the community change process
2. Evaluations of Complex Community Initiatives (CCIs) need a strong focus on the processes of community change
3. Evaluations of CCIs need to measure ongoing progress towards achieving outcomes and results in order to help a community guide its change process and hold itself accountable
4. Evaluations of CCIs need to understand, document, and explain the multiple theories of change at work over time
5. Evaluations of CCIs need to prioritize real-time learning and the community's capacity to understand and use data from evaluations

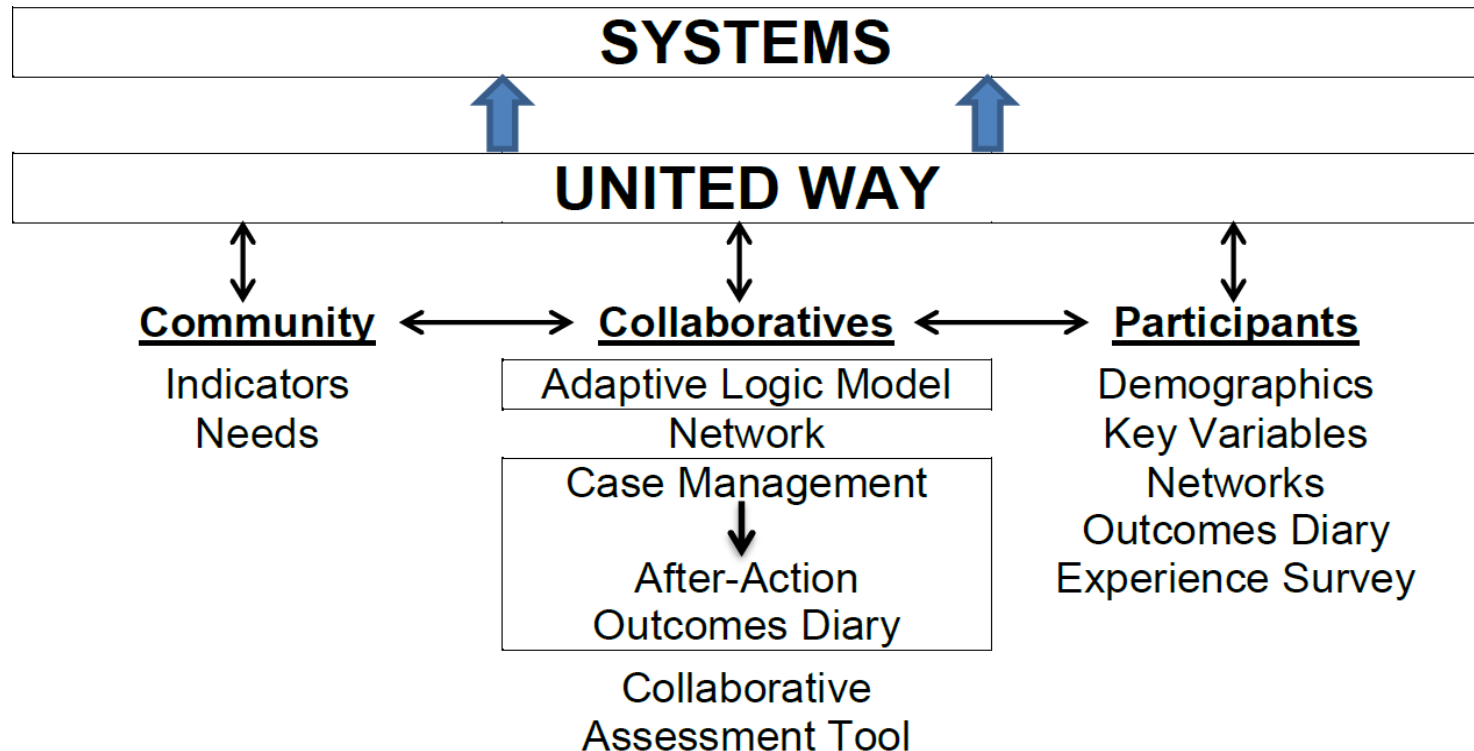
## DE Tool #1: Assessment tool for checking the innovation conditions

The following is a set of questions that organizations can ask themselves to see if they are in an appropriate space to apply developmental evaluation.

Question	Rationale
What is driving the innovation?	Developmental evaluation is particularly appropriate if an organization expects to develop and modify a program over the long term because of constantly shifting needs and/or contexts. It is helpful to discern between innovation taking place within an organization and the adoption of an external innovation, which may not need a developmental evaluation.
Are the proposed changes and innovations aimed at deep and sustained change?	Developmental evaluation is aimed at innovations that are driving towards transformational changes. Organizations often fine-tune their programs, and having an evaluative lens on those changes can be helpful; however the intensity of developmental evaluation may not be warranted in every instance.
Do we have a collaborative relationship with another organization in which there is innovative potential in combining our respective talents?	Developmental evaluation may help different organizations work together through the effort to innovate. In this situation, the developmental evaluator can help the organizations through some of the inevitable tensions of collaborating and can provide a measure of transparency about the experiment.
Under what conditions does the organization currently innovate? Is innovation part of the culture of the organization?	If this is already part of the culture, then the developmental evaluation role may be one that people within the team already play. If there is not a culture of innovation but there is a commitment to build one, then developmental evaluation may be helpful in stimulating that.
What are some core elements of what we do that we don't want to change?	There may be elements of an initiative that are known to work, or for another reason are expected to stay the same. Evaluation requires resources, and if things will not change, these resources are better directed elsewhere. If something is not going to be adapted but there is interest in finding out if it works or not, a summative evaluation is appropriate.
Is it clear for whom the evaluation is intended?	This is a vital question for any evaluation, developmental or otherwise. For an organization to make good use of developmental evaluation, it is important to have key decision makers interested in and open to using evaluative feedback to shape future actions. If the only user of the evaluation is external to the innovating team (such as a funder), then developmental evaluation is probably not the appropriate approach.

Source: Gamble. (2008) *A Developmental Evaluation Primer*. The J.W. McConnell Family Foundation

# Developmental Evaluation Example



# Types of Evaluation

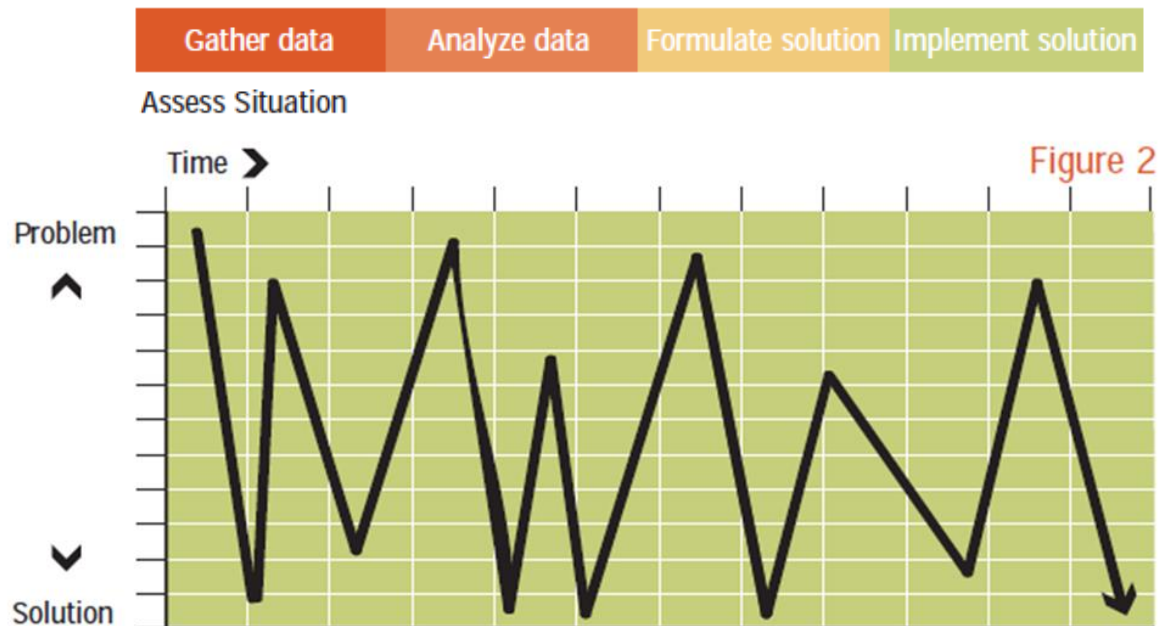
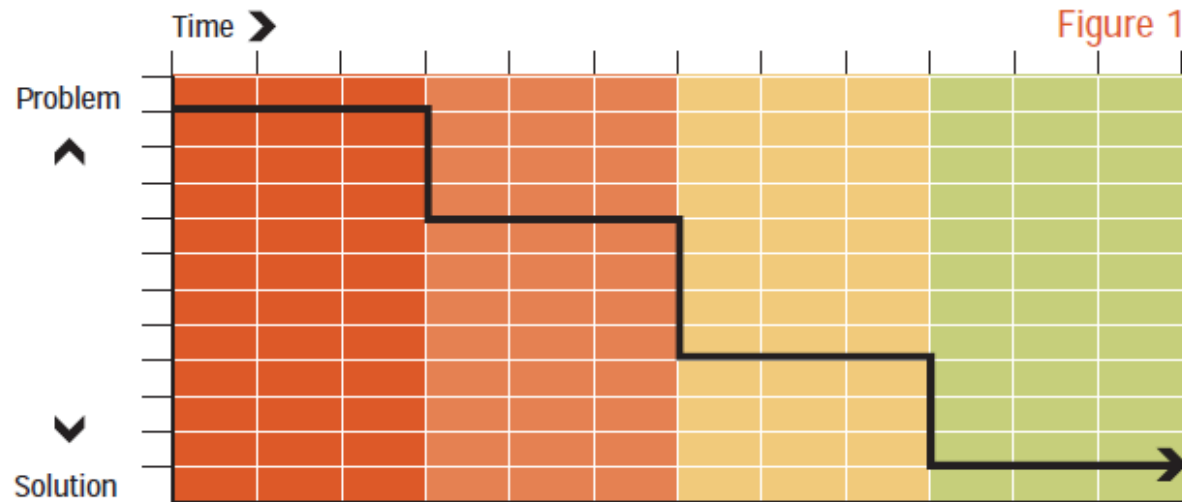
	Situation <sup>8</sup>
Summative evaluation	At the end of a program or initiative when key decisions about its future are going to be made. When judging the model's merit or worth for continuation, expansion, going to scale, or other major decisions.
Formative evaluation	When fine-tuning a model. When a future summative evaluation is expected and baseline data will likely be needed.
Developmental evaluation	When working in situations of high complexity. When working on early stage social innovations.

# Developmental Evaluation

Evaluation processes and activities that support program, project, personnel and/or organizational **development**. The **evaluator is part of the team** whose members collaborate to conceptualize, design, and test new approaches in a long-term, **on-going process of continuous improvement, adaptation, and intentional change**. The evaluator's primary function in the team is to elucidate team discussions with evaluative data and logic, and to facilitate data-based decision-making in the developmental process.

Michael Quinn Patton, Developmental Evaluation

# What is Developmental Evaluation?



Traditional evaluations	Developmental evaluations
Render definitive judgments of success or failure.	Provide feedback, generate learnings, support changes in direction.
Measure success against predetermined goals.	Develop new measures and monitoring mechanisms as goals emerge and evolve.
Position the evaluator outside to assure independence and objectivity.	Position evaluation as internal, team function integrated into action and ongoing interpretive processes.
Design the evaluation based on linear cause-and-effect logic models.	Design the evaluation to capture system dynamics, interdependencies, models and emergent interconnections.
Aim to produce generalizable findings across time and space.	Aim to produce context-specific understandings that inform ongoing innovation.
Accountability focused on and directed to external authorities, stakeholders and funders.	Accountability centered on the innovators' deep sense of fundamental values and commitment.
Accountability to control and locate responsibility.	Learning to respond to lack of control and stay in touch with what's unfolding and thereby respond strategically.
Evaluator determines the design based on the evaluator's perspective about what is important. The evaluator controls the evaluation.	Evaluator collaborates with those engaged in the change effort to design an evaluation process that matches philosophically with an organization's principles and objectives.
Evaluation results in opinion of success or failure, which creates anxiety in those evaluated.	Evaluation supports ongoing learning.

Source: Gamble. (2008) *A Developmental Evaluation Primer*. The J.W. McConnell Family Foundation

# Applying Developmental Evaluation

- **Accompaniment**
- **Collecting Data**
- **Framing and Reporting**
  - There is a distinction between the notes of a developmental evaluator and the more traditional minutes of a meeting. The developmental notes may identify:
    - process observations;
    - points of tension;
    - implicit decisions;
    - assumptions made; or
    - emerging themes and patterns.
- **Strategy**
  - What evidence would indicate that the process is working? Or not working?
  - What are the organization's real-time feedback mechanisms for tracking changes/growth?
  - What could go wrong and how would we know? And when things go right, how do we know that? Why were we successful? How can we learn from our success?
  - Given where we are (baseline) and where we want to go, what are the foreseeable decision points and timelines at which we determine how we're doing? What information will we want at those decision points to make any needed adjustments?
- **Indicators**



# Developmental Evaluation Tools

- **What? So what? Now what?**
  - One of the basic frameworks for evaluation, aimed at simplifying what we do, is summarized by asking three questions: What? So what? Now what? These simple questions help us to analyze multiple factors and to align diverse questions and actions towards common interests.
    - **What?** What do we see? What does data tell us? What are the indicators of change or stability? What cues can we capture to see changing patterns as they emerge?
    - **So what?** So, what sense can we make of emerging data? What does it mean to us in this moment and in the future? What effect are current changes likely to have on us, our clients, our extended network and our field of inquiry and action?
    - **Now what?** What are our options? What are our resources? When and how can we act - individually or collectively – to optimize opportunities in this moment and the next?

# Developmental Evaluation Tools, Cont.

- Network mapping
- Revised and emergent modeling
- Simulations and rapid reconnaissance
- [Appreciative inquiry](#)
- Visual language
  - [http://www.visual-literacy.org/periodic\\_table/periodic\\_table.html](http://www.visual-literacy.org/periodic_table/periodic_table.html)
- After-Action Report

# Evaluating Network Connectivity

Pillar	Focus	Example Evaluation Questions	Casebook Examples (evaluation funder)
1) Connectivity	<b>(a) Membership</b> <i>The people or organizations that participate in a network</i>	<ul style="list-style-type: none"> <li>• Who participates in the network and what role does each member play?</li> <li>• Who is connected to whom? Who is not connected but should be?</li> <li>• Has the network assembled members with the capacities needed to meet network goals (experience, skills, connections)?</li> <li>• Is membership adjusted to meet changing network needs?</li> </ul>	<ul style="list-style-type: none"> <li>• Reboot (<i>Jim Joseph Foundation</i>)</li> </ul>
	<b>(b) Structure</b> <i>How connections between members are structured and what flows through those connections</i>	<ul style="list-style-type: none"> <li>• What are the number, quality, and configuration of network ties?</li> <li>• What is flowing through the network — information and other resources?</li> <li>• How efficient are the connections the network makes?</li> <li>• How dependent is the network on a small number of individuals?</li> <li>• Is structure adjusted to meet changing network needs and priorities?</li> </ul>	<ul style="list-style-type: none"> <li>• Barr Fellows Program (<i>Barr Foundation</i>)</li> <li>• Safe Schools/Healthy Students (<i>federal government</i>)</li> </ul>

# Evaluating Network Health

2) Health	<b>(a) Resources</b> <i>The material resources a network needs to sustain itself (e.g., external funding)</i>	<ul style="list-style-type: none"> <li>• Has the network secured needed material resources?</li> <li>• What type and level of resources does the network have?</li> <li>• How diverse and dependable are these resources?</li> <li>• How are members contributing resources to the network?</li> <li>• Is the network adapting its business plan over time?</li> </ul>	<ul style="list-style-type: none"> <li>• Urban Sustainability Directors Network (<i>multiple funders</i>)</li> <li>• RE-AMP (<i>Garfield Foundation</i>)</li> <li>• KnowHow2Go (<i>Lumina Foundation</i>)</li> </ul>
	<b>(b) Infrastructure</b> <i>Internal systems and structures that support the network (e.g., communication, rules and processes)</i>	<ul style="list-style-type: none"> <li>• What infrastructure is in place for network coordination and communications?</li> <li>• Are these systems efficient and effective?</li> <li>• What are the network's governance rules and how are they followed?</li> <li>• Do decision-making processes encourage members to contribute and collaborate?</li> <li>• How are the network's internal systems and structures adapting?</li> </ul>	
	<b>(c) Advantage</b> <i>The network's capacity for joint value creation</i>	<ul style="list-style-type: none"> <li>• Do all members share a common purpose for the network?</li> <li>• Are members working together to achieve shared goals, including goals that emerge over time?</li> <li>• Are all members contributing to network efforts?</li> <li>• How are members adding value to one another's work?</li> <li>• Are members achieving more together than they could alone?</li> </ul>	

# Evaluating Network Results

## 3) Results

### (a) Interim Outcomes

*Results achieved as the network works toward its goal or intended impact*

- Are there clear signals of progress/interim outcomes for the network and are they understood and measured by members?
- Is the network making progress on interim outcomes that signal progress on the way to longer-term goals or intended impacts?

- MA Regional Networks to End Homelessness (*Paul and Phyllis Fireman Charitable Foundation*)

### (b) Goal or Impacts

*The ultimate goal or results the network is after*

- At which level(s) are impacts expected — on individual members, on members' local environments, and/or on members' combined impact on their broader environment?
- If the goal is achieved or ultimate impacts observed, can a plausible and defensible case be made that the network contributed to them?

- The Fire Learning Network
- International Land Coalition

# After-Action Reports



## After Action Reviews and Retrospects

This guide has been developed using materials provided by the Overseas Development Institute's Research and Policy in Development (RAPID) Programme



### WHAT IS AN AAR?

Organizational learning requires a continuous assessment of organizational performance, looking at successes and failures. This ensures that learning takes place and supports continuous improvement. The After Action Review (AAR) is a simple tool that facilitates this assessment.

It works by bringing together a team to discuss a task, event, activity or project, in an open and honest way.

### THE PROCESS

There are many different ways to conduct AARs. The whole process should be kept as simple and as easy to remember as possible. The essence of an AAR is to bring together the relevant group to think about a project, activity, event or task, and reflect on the following simple questions:

- What was supposed to happen?
- What actually happened?
- Why were there differences?
- What worked?
- What didn't?
- Why?
- What would you do differently next time?

A **Retrospect** has a similar format to an After Action Review, but asks the following more detailed questions:

- What was the objective of the project?
- What did we achieve?
- What were the successes? Why? How can we repeat the success?
- What was disappointing? Why? How can we avoid them in future?
- 'Marks out of 100', what would move it closer to 100?

### BENEFITS

The After Action Review is a powerful tool because it produces quick results in a short time and can be applied to a broad range of activities.

Its strengths are the following:

- it allows team members to immediately apply lessons learned;
- it gathers the group's intuitions about the strengths and weaknesses of an activity or a project;
- it gives team members an opportunity to share their views and ideas and thus develop a common perspective on which they can base their future work.

### KEY POINTS AND PRACTICAL TIPS

- ✓ Post the questions on flipchart sheets prior to the session. Write answers on the sheet as the session progresses.
- ✓ The facilitator should prepare some lead-in questions and may have to directly solicit answers.
- ✓ If there are issues with either openness or time, it may be worthwhile to gather individual ideas first and then facilitate a group discussion.
- ✓ An uninvolved note-taker should be asked to take minutes for the session. This will make sure lessons learned are captured.
- ✓ Actionable recommendations should be as specific as possible. For example, an AAR could have the following recommendation: 'Make contact with the organizing body representative and ask about the range of participants before planning the workshop.'
- ✓ Participants of an AAR should include all members of the team.
- ✓ AARs should be carried out immediately, while the team is still available and memories are fresh.

### 1. Framing the issue

Social innovators are mobilized by a powerful sense that something needs to change. They may have a new perspective or approach to a historically stubborn issue, or may see, in a new way, the intersection between multiple issues. As innovators work on these issues, their understanding moves from a vague understanding to increased clarity. New learning may cause a shift in thinking which prompts another cycle of uncertainty and clarification. Developmental evaluation supports innovators in the conceptualization and articulation of the problem, by helping to frame the issue and its dynamics.

### 2. Testing quick iterations

Many people who develop and deliver social programs naturally experiment.<sup>10</sup> New ways of doing something are tried, often based on feedback loops and perspective about changing needs and demands, which can lead to improvements. Developmental evaluation brings a measure of rigour to the learning generated from these experiments. As new programs roll out, leaders intuitively make observations and refinements. These lessons are usually part of what is our natural private learning processes. Developmental evaluation is intended to make visible the intuitive and the tacit. Applying developmental evaluation means being more systematic about subjecting relevant data and observations to interpretation and judgment.

### 3. Tracking the trajectory of the innovation

A standard characteristic of problem solving is that once the problem solver experiences the "eureka moment," the path to the solution seems obvious. When innovators look at projects retrospectively, the description of going from beginning to end appears seamless and direct. Key insights about how something was successfully accomplished are often inaccessible, which doesn't help the next person trying to solve a similar problem, or the original innovator in trying to apply the learning process in other situations. Developmental evaluation records the roads not taken, unintended consequences, incremental adjustments, tensions and sudden opportunities. The tracking reveals what it takes to create something new, which serves two purposes: 1) it makes the decisionmaking along this path more transparent and 2) it generates valuable data useful for dissemination. Such documentation also supports accountability while allowing for a high degree of flexibility.

# Myths about Developmental Evaluation

- **Myth #1: Developmental evaluation replaces other evaluations**
  - Developmental evaluation is not appropriate to all situations. It is not superior, or inferior, to formative and summative evaluation. Rather, DE is an addition to the current set of evaluation approaches. Deciding when to do various evaluations – summative, formative or developmental – should be a purposeful decision.
- **Myth #2: Developmental evaluation is about soft methods**
  - Developmental evaluation is as rigorous as any evaluative process. Like all good evaluations, it is evidence-based.
- **Myth #3: Developmental evaluation is about collecting stories**
  - Story collecting may be used, but this also occurs with several other data collection processes. Developmental evaluation may involve qualitative or quantitative methods, or both.



# Myths about Developmental Evaluation, Cont.

- **Myth #4: Developmental evaluation is process evaluation**
  - The ultimate focus is results. Process is attended to, but developments that move something towards outcomes is the ultimate objective. Outcome information is not counter to developmental evaluation; in fact it very much informs it. Using a developmental evaluation approach invigorates interest in generating data on outcomes and in working through reasoned processes to gather and interpret it.
- **Myth #5: Developmental evaluation downplays accountability**
  - The accountability of developmental evaluation rests in its ability to support development. If nothing is developed, it has failed. Learning what does and doesn't work is a type of development. Deeper questions may be a developmental result, but something must be developed.
- **Myth #6: Developmental evaluation is the same as participatory evaluation**
  - Participatory evaluation is about a distinction in approach, where developmental evaluation is about a distinction in purpose. Participatory approaches can be used to inform summative, formative and developmental evaluations. Developmental evaluation is particularly oriented to supporting early stage innovations in complex environments. A participatory approach makes a lot of sense in developmental evaluation because of the need for high trust and quick feedback.

# Group Exercise

## Developmental Evaluation Framework

- Diagram a developmental evaluation strategy for the Allegheny County initiative.

# Evaluative Thinking (ET)

- **Create an intentional ET learning environment**
  - Display logic models in the workplace
  - Create public spaces to share questions and ideas
  - Highlight learnings, both good and bad as they appear
  - Have teams create logic models together
- **Schedule regular meetings to focus on ET practice**
  - Mine logic models for information about assumptions
  - Encourage questions such as, what plausible explanations are there for this finding?
  - Take on various stakeholder perspectives (role play) to think about program activities
  - Diagram or illustrate thinking with colleagues