SESRI
Policy & Program Evaluation Workshop

Doha, Qatar
January 19-22, 2015

Outline: Session 1

- **Workshop objectives**

- **Introductions**

- **Creating public programs to address public problems**

- **Defining program goals (outcomes), targets, instruments (inputs), and results (outputs)**

- **Using program models to define a theory of action**

- **Developing hypotheses from theories of action**

By the end of this workshop, you should be able to:

- Understand the purpose of evaluation in public policy

- Identify the primary components of policy and program evaluation

- Consume evaluation research in an informed, systematic manner

- Consider the trade-offs inherent in designing evaluations

Who are we?

- **Michael Traugott (Mike)**

- **Elisabeth Gerber (Liz)**

- **Ann Lin**

- **Monica Bhatt**

- **Fatimah Alkhaldi**

What constitutes a public policy *problem*?

- A problem affecting some segment of society that government action could (but may or may not) address

- Potential government actions include proclamations, decrees, informal policy, lack of policy ("non-policy")

- Example of climate change in Doha:
  - Officials cannot solve changing weather patterns, which is the root of the problem
  - However, officials can address the problems that arise as a result (e.g., flooding)

Traffic in Qatar

Qatar Tribune,

April 8, 2014

A million accidents cost Qatar QR18 bn in 6 years

- 1,289 deaths between 2007 & 2012, according to a QRSSC report
- February 22 Street a prominent 'black spot'

TRIBUNE NEWS NETWORK
DOHA

A study conducted by the Qatar Road Safety Studies Center (QRSSC) at Qatar University has found that Qatar lost a staggering QR17.6 billion as a result of road accidents between 2007 and 2012. Qatar recorded a total of 1,289 fatal road accidents and a million mishaps, ranging between serious and minor, during the six-year period.

The study - carried out by a group of students from College of Engineering at Qatar University - is trying to compute the cost of traffic accidents and the consequent damage to both the individual and society in general.

The report categorised 'black spots' as places notorious for traffic bottlenecks and which frequently see accidents.

The researchers found that traffic accidents in 2007 alone cost Qatar about QR2.5 billion. The amount soared to QR3.2 billion by 2012.

The study noted that Qatar spent an average of approximately 2.7 percent of the GDP on traffic accidents annually. It said that since Qatar is considered a developing country, the cost of traffic accidents has become equivalent to about 2 to 5 percent of the national income.

According to the study, among the 1,289 recorded fatalities, 523 were of drivers, 377 of passengers, and 389 pedestrians. Of them, 1,197 were males and 92 females, the research reveals.

The researchers have also noted about the problem of 'black spots'. The report categorised 'black spots' as places notorious for traffic bottlenecks and which frequently see accidents. The QRSSC is working on a computer simulation of the February 22 Street - a prominent 'black spot' in Qatar known for its traffic jams and accidents.

What makes a public problem "public"?

- **Public goods**

- **Societal needs**

- **Public perception**

- **Political pressure**

- **Concerns about values**

- **Others?**

What is a program?

- **Instruments (inputs) used to achieve a policy goal (outcomes)**

- **Bounded by time, scope or population**

- **Evaluation requires specific goals, targets, inputs, and outputs**

- **Example**
  - **Paying teachers to teach in rural areas (input), in order to improve rural education (goal), as measured by the test scores (output) of rural schoolchildren (targets)**

Clicker Question 1

A program that addresses traffic congestion in Doha should … (Click what you think the government's goal is.)

a) Reduce the number of traffic accidents, in order to improve the health and lower the mortality rate.

b) Reduce air pollution, caused by idling vehicles and under-utilization of carpools and mass transit.

c) Reduce travel times, in order to increase business productivity and quality of life.

d) All of the above.

Programs require . . .

- **GOALS/OUTCOMES**
  - **What the policy hopes to achieve**

- **TARGETS**
  - **People and organizations slated for change**

- **INSTRUMENTS/INPUTS/INTERVENTIONS**
  - **Mechanism by which change happens**

- **OUTPUTS**
  - **Change that is slated to occur**

Goals

- What does the policy hope to achieve?
- Are there multiple goals?
- What are the tensions among them?
- What are the assumptions inherent in these goals?

Clicker Question 1 (again)

A program that addresses traffic congestion in Doha should… (Click the ONE that **you** think should be the goal.)

a) Reduce the number of traffic accidents, in order to improve the health and lower the mortality rate.

b) Reduce air pollution, caused by idling vehicles and under-utilization of carpools and mass transit.

c) Reduce travel times, in order to increase business productivity and quality of life.

Targets

- **Which individuals or groups is the policy designed to affect?**
- **Who are the recipients of the program?**
- **How are they chosen?**
- **Who delivers the program?**

Exercise

Turn to your neighbor. Who are the right target(s) for a program with the goal that we chose in the previous clicker question?

<u>Possibilities</u>:

- Drivers: Commuters, commercial drivers, reckless drivers

- Businesses: Mass transit operators, companies with workers who can telecommute, companies who get deliveries

- Service providers: Driving instructors, schools

Inputs

- **Also called program instruments, program interventions, program treatments**

- **Can be rules, education, incentives, sanctions, opportunities, infrastructure**

- **Must be linked to outputs**

Exercise

- **Turn to your neighbor. Propose an input to reduce traffic congestion that would be appropriate for the following targets:**
  - **Bad drivers**
  - **Owners of businesses with workers who could telecommute (who could work from home)**
  - **People living in residential neighborhoods located near major traffic routes**

Outputs

- **Also called program results**

- **Must be subject to <u>change</u> and <u>assessment</u>**

- **Can be anticipated or unanticipated**

- **Different from program outcomes or goals:**
  - The evaluator should choose outputs that have the closest connection possible to the program inputs.
  - Outputs indicate outcomes, but are not equal to them; evaluators should be skeptical of the output-outcome relationship.
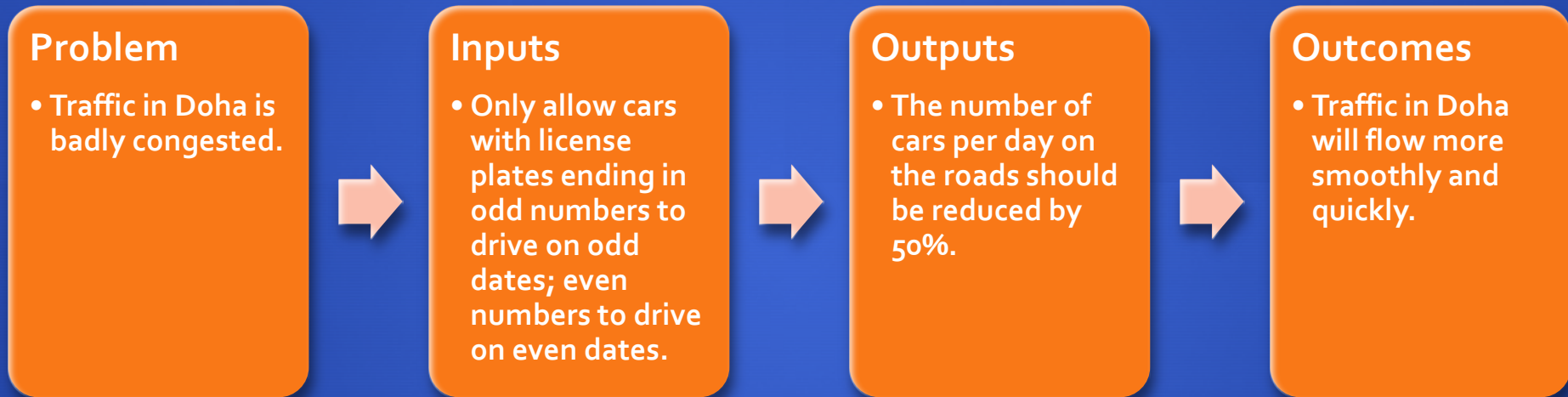
Clicker Question 2

Which of these pairs connects an input with an appropriate output?

a) Fining drivers who cause accidents   -> more money collected in fines

b) Fining drivers who cause accidents -> fewer accidents

c) Fining drivers who cause accidents - > fewer traffic jams

d) All of the above.

Program Model

- Begins with an understanding of what a program is supposed to produce, and for whom -- e.g., goals and targets

- Transforms that understanding into a mechanism – specific inputs and outputs

- Displays, often symbolically, a "theory of action" (how the program *is supposed to work*)

- Implies causal relationships between elements of the model ("If this…then this…"

- Also called: logic model, conceptual map, etc.

The Simplest Form

Problem
- Traffic in Doha is badly congested.

→

Inputs
- Only allow cars with license plates ending in odd numbers to drive on odd dates; even numbers to drive on even dates.

→

Outputs
- The number of cars per day on the roads should be reduced by 50%.

→

Outcomes
- Traffic in Doha will flow more smoothly and quickly.

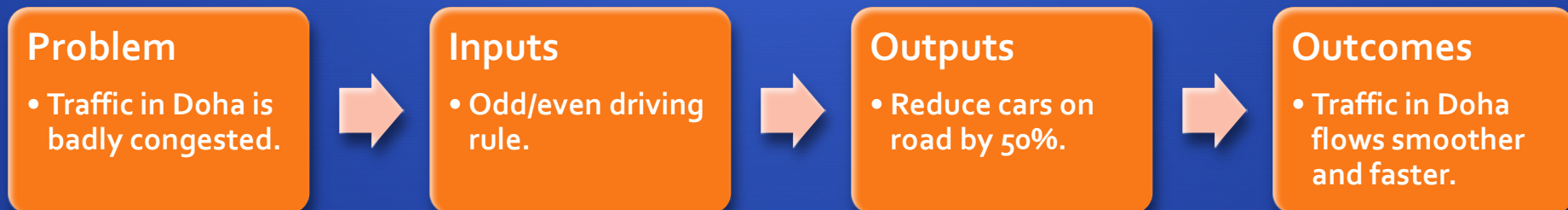**What is the causal story? (What "causes" congestion?)**

The Importance of Assumptions

- Includes the beliefs we have about the program, its participants, or how it might work

- May or may not be stated explicitly

- Typically not tested

- Program models can help make these assumptions explicit, but not always

- Evaluator must be aware of what assumptions are inherent in the program model

Clicker Question 3

What assumptions are embedded in the odd/even driving rule?

a) Congestion is due to too many cars on the roads, rather than to inefficient road design.

b) Drivers have only one car per driver.

c) Drivers are unable to get waivers from the odd/even rule.

d) Drivers will not take advantage of newly empty streets to idle or park their cars illegally.

Problem
- Traffic in Doha is badly congested.

➡

Inputs
- Odd/even driving rule.

➡

Outputs
- Reduce cars on road by 50%.

➡

Outcomes
- Traffic in Doha flows smoother and faster.

Assumptions can be about…

- **Program staff – knowledge, skills, will**

- **Available resources**

- **Target motivation and behavioral patterns**

- **Causal links between elements of the program model**

- **External environment**

- **Extant knowledge base**

Hypotheses

- **Specifies expected relationship between elements of the program that will be tested with data**

- **Differ from assumptions which are not tested, but which are important to clarify when testing hypotheses and evaluating a program**

- **Program evaluators use hypotheses in conjunction with data to test the relationship between elements of program model**

Clicker Question 4

Which is a testable hypothesis that may be formulated based on the given program model?

a) Odd/even driving rules are the best method of reducing traffic congestion.

b) A reduction in cars on Doha roads will reduce traffic congestion.

c) Are individuals who comply with the odd/even rule law-abiding?

d) Traffic congestion in Doha is caused mainly by rude drivers.

Problem
- Traffic in Doha is badly congested.

Inputs
- Odd/even driving rule.

Outputs
- Reduce cars on road by 50%.

Outcomes
- Traffic in Doha flows smoother and faster.

SESRI

Policy & Program Evaluation Workshop

Doha, Qatar

January 19-22, 2015

Outline: Session 2

- **What is causality?**

- **Why is causality important for program evaluation?**

- **Randomized Control Trials (RCT)**

- **Case study: Qatar Financial Literacy Study**

- **Group exercise**

What is causality?

- **A directional link between two events, i.e., cause and effect. A➜B**

- **To exhibit a causal relationship, two variables must satisfy three conditions:**
  - Statistical relationship: A and B covary
  - Temporal priority: A precedes B
  - Alternative explanations eliminated: C is not the real cause of B

- **Establishing causality is key for program evaluation**
  - The program model implies causality
  - However, it is difficult to establish
  - *ESTABLISHING CAUSALITY IN PROGRAM EVALUATION IS A MATTER OF DESIGN.*

Clicker Exercise – Causal Hypotheses

- Which of the following are causal hypotheses? Click all that apply.
  a) Migrant workers with higher levels of education save more of their wages.
  b) Students prefer interactive classroom pedagogy over lecture-based pedagogy.
  c) Qatarization of teaching faculty is too expensive.
  d) Replacing roundabouts with traffic lights reduces traffic accidents.

Why do we care if A → B?

- Shows whether a program "works" or not

- May explain why a program works (often called the "mechanism")

- Useful for making predictions

- Can aid in expanding or replicating program

Counterfactuals

- In order to establish a causal relationship, we need a *counterfactual*, or an alternative situation in which everything is *exactly identical* except that the participants do not experience the *treatment* or *intervention* (the program input).

- This is impossible! However, we can use experiments and other research designs to approximate this counterfactual.

The Role of Experiments

- **One way to establish a strong counterfactual is through the use of experiments:**

Group	Treatment Received?	Outcome
Group A (Treatment Group)	✔	Treatment Effect
Group B (Control Group)		Outcome for "Business As Usual"

- **Most rigorous research design is a Randomized Control Trial (RCT)**

Randomized Control Trials (RCTs)

- Powerful research design in which the researcher/evaluator controls assignment of the treatment.

- RCTs rely on *random assignment* to create a counterfactual
  - Researcher randomly assigns individuals in a study to two groups:
    - Treatment
    - Control
  - Each individual must have an equal chance of being assigned to either group

- This creates *groups* that are "equal in expectation" *even if the individuals are not identical.*

Why does random assignment work?

- **Ensures that the groups are equivalent (at least in expectation of receipt of treatment) prior to being treated or not**
    - **This provides a defensible *counterfactual*, which then allows us to establish *causality***
    - **Creates "all else equal" conditions across two groups**

- **Allows researcher to know and control the selection process correctly**

- **Ensures alternative causes are not confounded with participation in the program**

Group Exercise

- Imagine we formed two groups in this room - a front half and a back half - to work on a group project.
  - Are the two equal in terms of major characteristics?
    - Will the two have equal numbers of men and women? Qataris and ex-pats? Arabic and English speakers? Other characteristics?
  - Are they "equal in expectation"?
    - Why is this a different question than the ones above?

- Now, imagine we formed two groups in this room using random assignment.
  - Are the two equal in terms of major characteristics?
  - Are they "equal in expectation"?

Examples of RCTs

Example: Savings Intervention

- **Does education increase the savings rate among low-income migrant workers?**

    - **Why might the lack of migrant savings constitute a *public policy problem*?**

    - **What type of program or policy might be designed to address low savings rates among migrant workers in Qatar?**

    - **What kinds of populations should such a program target? How would participants be recruited? What implications might that have for program evaluation design?**

Considerations for Program Evaluators

- **How can we know for sure whether the program actually changes savings behavior (*internal validity*)?**
  - Individuals who choose to participate in the program may be systematically different than those who do not (*selection bias*).
  - This could make it look like participants change their behavior as a result of the program, while in fact they were predisposed to such behavior even in the absence of the program.

- **How can we know for sure whether the program targets the individuals we are interested in (*external validity*)?**
  - Who we recruit has implications for the claims we can make from the study results.

Group Exercise

- **Turn to your neighbor and discuss why individuals who participate in the program may be different than those who do not.**

- **What implications would these differences have for our evaluation design?**

Evaluation Design Options

- How should participants be recruited? How should the treatment and control groups be constructed?

- These choices determine our ability to construct a strong counterfactual
  - Option 1: compare savings behavior of individuals before and after they complete the program
  - Option 2: compare savings behavior of individuals who complete the program versus those who do not
  - Option 3: provide the program for those individuals who are vulnerable financially and compare outcomes
  - Option 4: provide the program for those individuals who are most likely to be responsive to the program and compare outcomes

"Motivating Migrants" study Seshan and Yang

- RCT to study financial habits of male migrant workers in Qatar and their wives who remained in Kerala, India (N=232). Average annual income was US $6,175. Average remittances sent home was $2,637. Average savings was $2,395.

- *Intervention*: Migrant workers were randomly invited to a workshop on financial savings that lasted a few hours for a single setting, provided in November 2010 (N=157).

- *Measurement*: Baseline survey sent out in 2010 prior to workshop, follow up survey sent out in 2012 to husbands and wives.

- *Results*: Wives of treated migrants changed their financial practices and were more likely to report seeking out financial education. Treated migrants and their wives more likely to make financial decisions jointly.

Program Model:
Motivating Migrants study

Problem
- Migrant worker savings rates are very low, leaving their families financially vulnerable.

Inputs
- Financial literacy program training sessions for both husbands and wives.

Outputs
- Wives are more involved in family financial decisions, lowest income families save more.

Outcomes
- Families are more financially secure.

Moving from Program Models to an RCT

Program Model	Evaluation Terminology	Example
Goals	Hypothesis	Increase savings
Program inputs	Intervention / Treatment	Financial literacy sessions
Program inputs	Independent Variables	Participating in FL sessions, demographic characteristics, other covariates
Outcomes	Dependent Variables	Savings rates, family financial decision-making
Targets	Treatment Group	Individuals who participated in the FL program
N/A	Control Group	Otherwise similar individuals who did not participate in FL program
Stakeholders	N/A	Broader population, government entities, employers

Discussion of Seshan & Yang Program

- **Strengths?**

- **Weaknesses?**

Further Considerations

- **The RCT helps to rule out confounding factors and attribute causality to the program inputs. But, we may want to know additional information to make sure.**
  - **What if members of the treatment group spoke to members of the control group and shared what they learned in the sessions (*contamination of treatment*)?**
  - **What if people who would be unwilling to change their savings behaviors did not volunteer to join the randomized study (*contamination of control*)?**
  - **What if participating in the study caused people to change their behavior in ways that are not measured by the study (*collateral outcomes*)?**
  - **What if the people selected to be in the treatment group did not end up participating in the intervention (*selection bias*)?**

Take Home Exercise

- In your packets, you will find a Day 1 Take Home Exercise. Please complete this tonight and bring a printed or hand-written copy with you tomorrow.

- You will submit your responses to the teaching assistants at the beginning of class.