

# Linking “M” to IE: the role of monitoring in impact evaluation

**Barbara Bruns**  
**Program Manager, SIEF**  
**and Lead Economist**  
**Human Development Network**

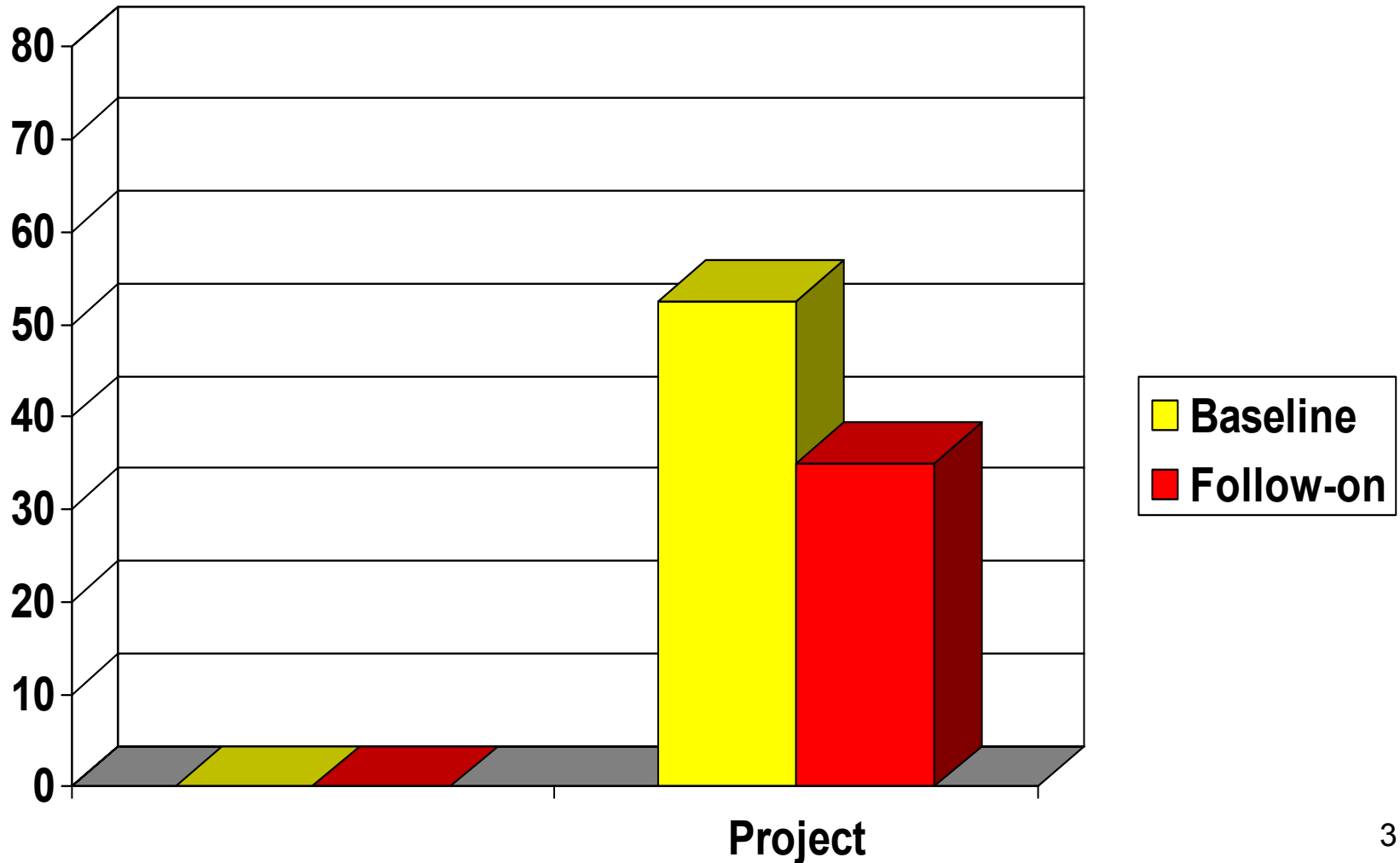
**Manila**

**December 2008**

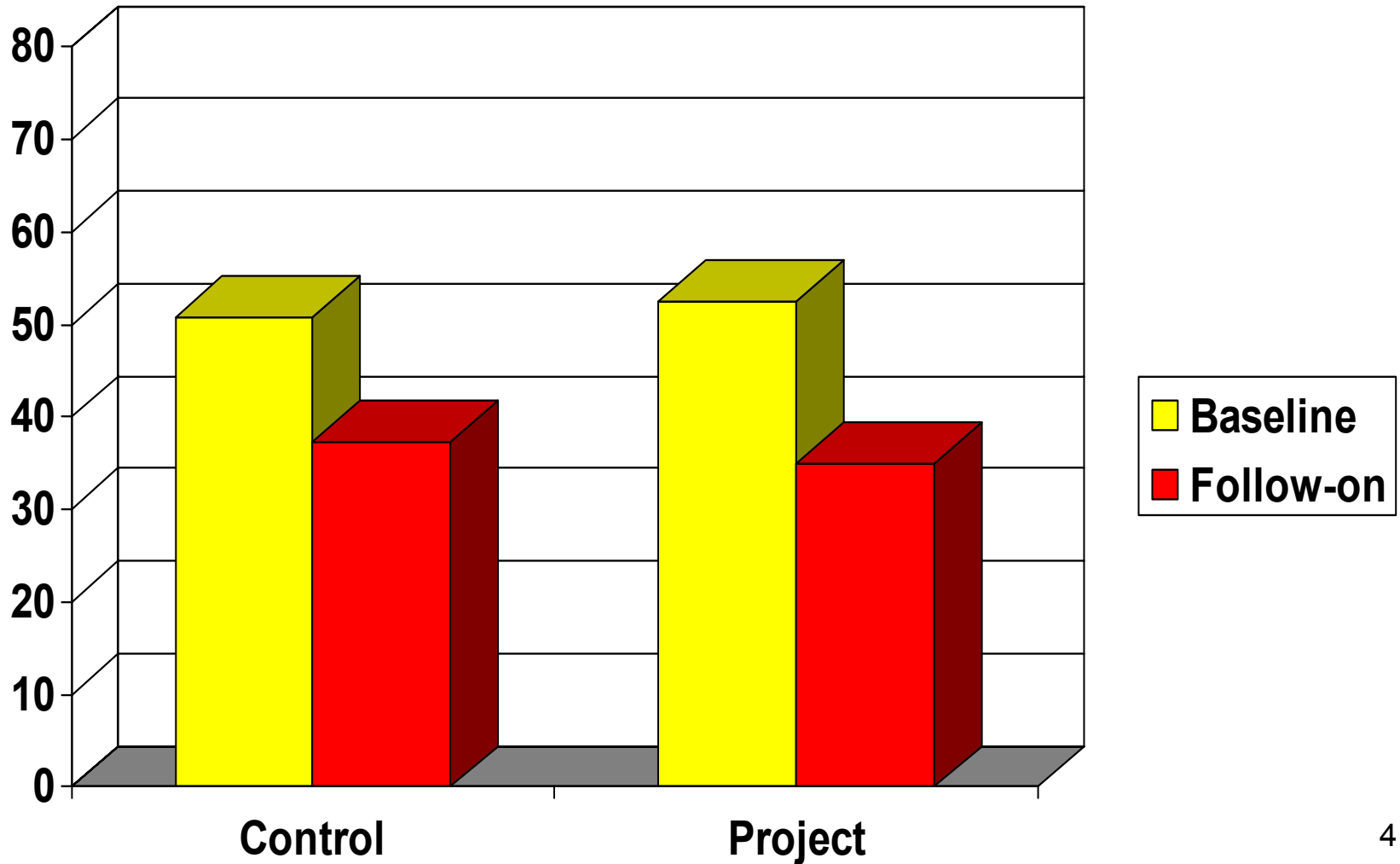
# Objectives of this session

1. Why monitor?
2. How to monitor
  - Using a RESULTS chain
3. How to measure what you want to know
  - Selecting indicators
  - Timing data collection
  - Mixed methods

# **% of Children <5 who were moderately or severely malnourished – World Bank supported nutrition program**

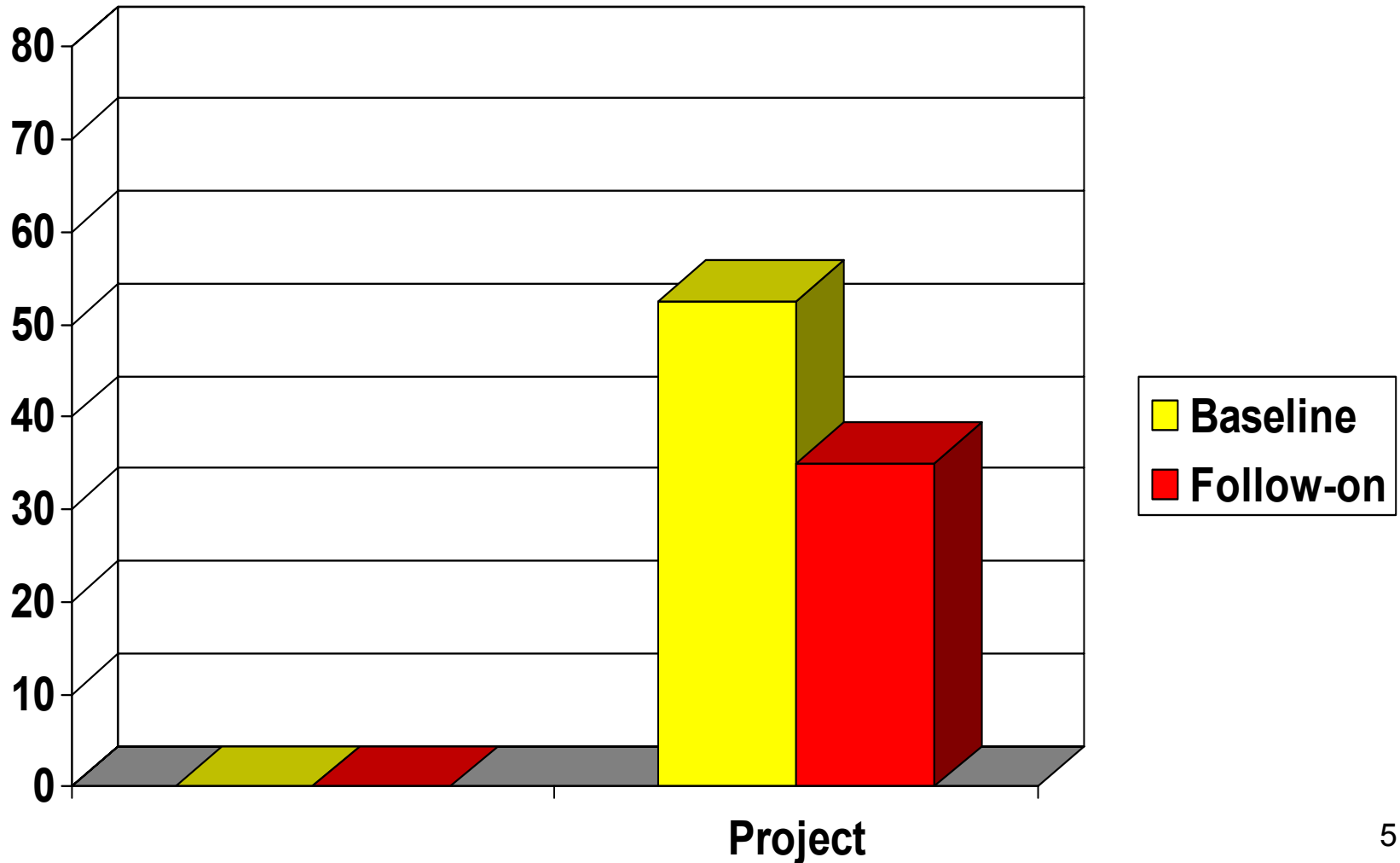


# ...compared to control group

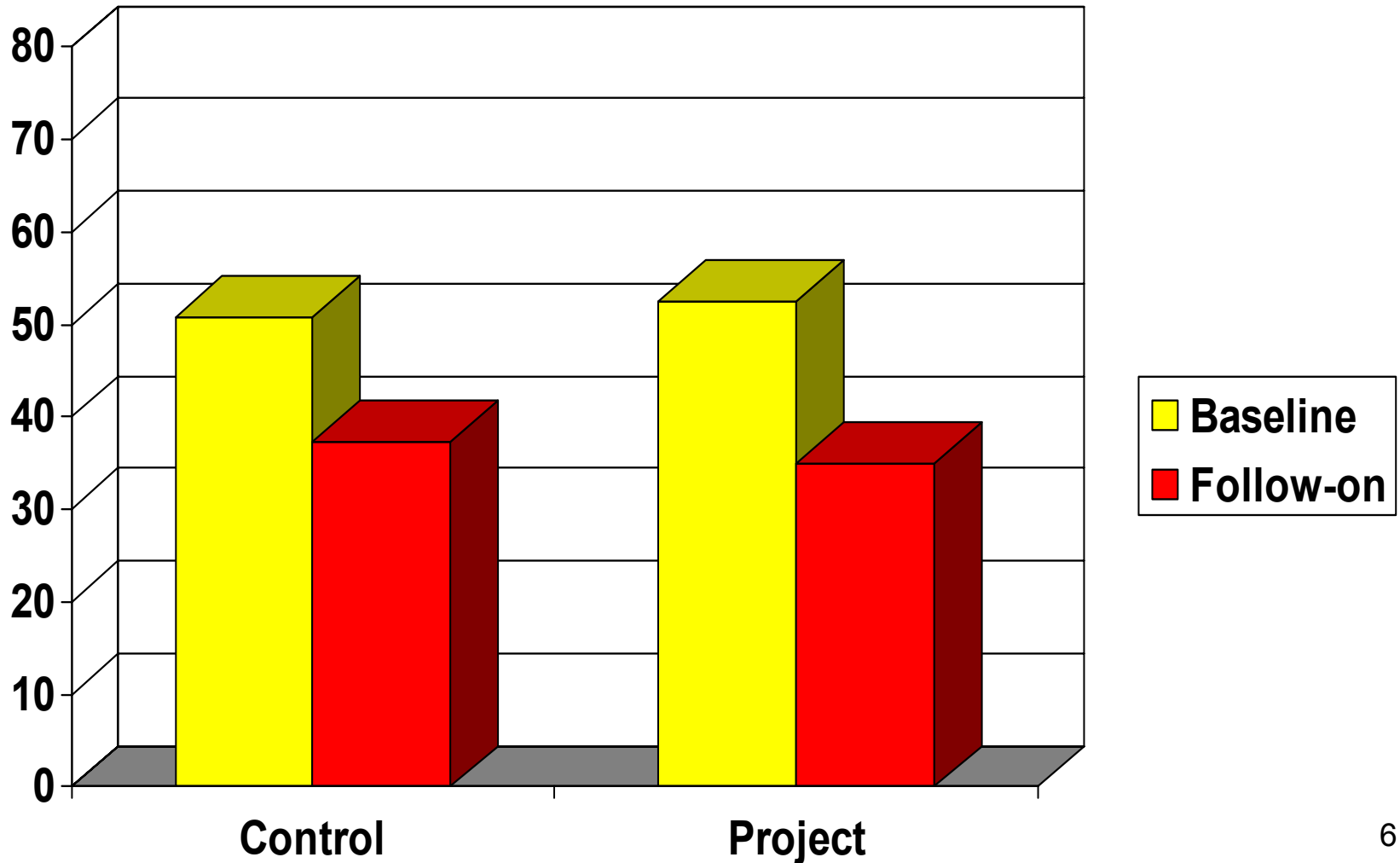


# But....feeding intervention arrived late –

only 50% of program districts reached by year 3



**...and also reached 25% of control communities**



# Mantras for this week

- To do good IE, you need to understand how the data were generated
- **...and how the program was implemented!**

## ...this is called Monitoring

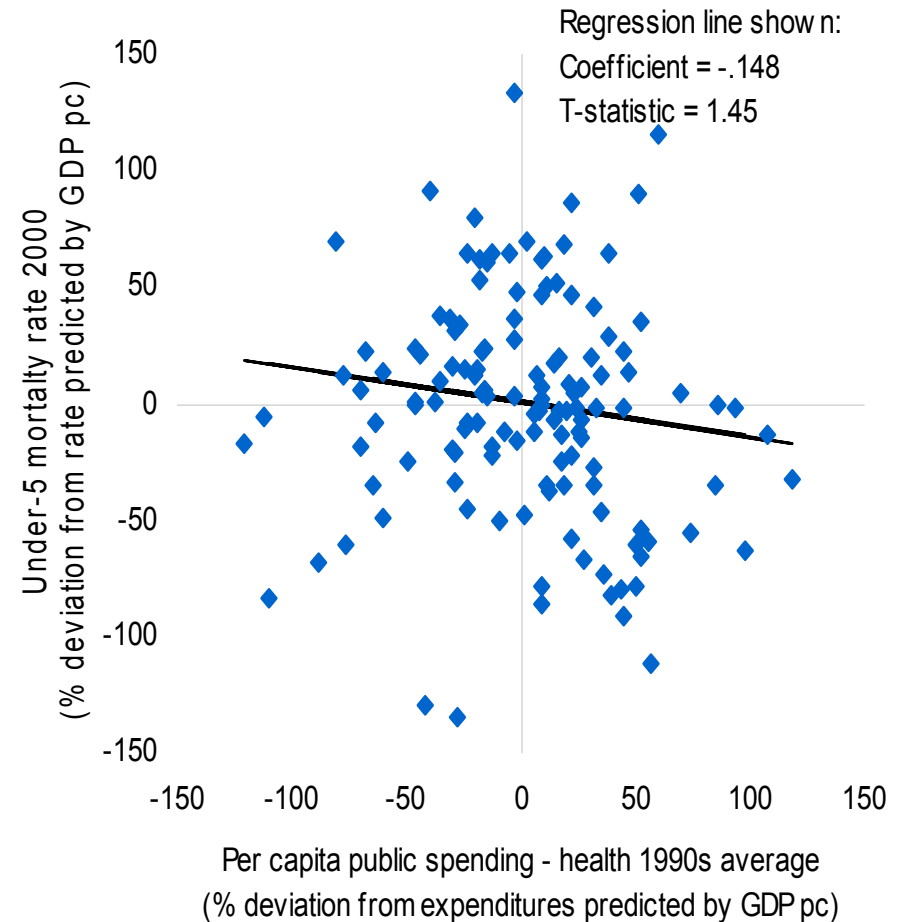
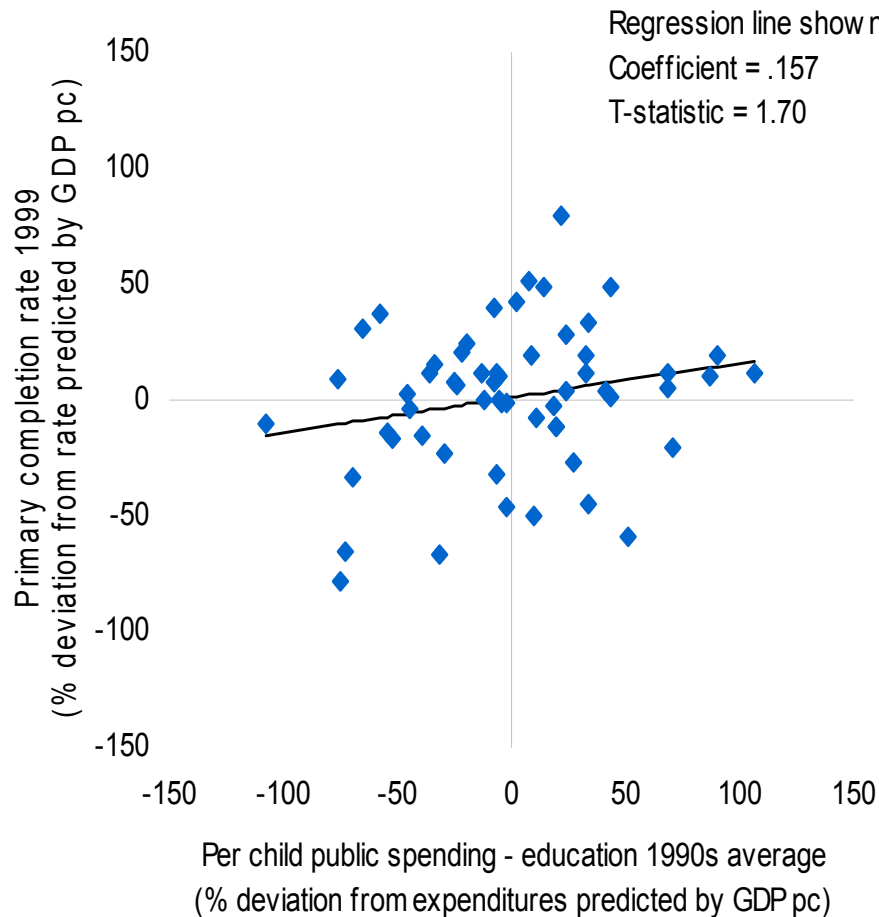
- Ongoing, real time data collection on inputs, outputs and outcomes to track implementation
  - Follows the transformation of inputs to outcomes
  - Identifies key indicators
  - Generates real time information on implementation process
  - Essential management/accountability tool
- A crucial part of impact evaluation design

# Essential function in business

<b>Decision Making Process</b>	<b>Private Enterprise</b>	<b>Social Programs</b>
1. Identifying Key Results	Increased shareholder value	Improved social welfare
2. Measuring Results	Revenues-Costs= Profits	Measure against counterfactual
3. Monitoring Process	Key to understanding what drives costs and profits	Often focused on inputs and outputs (which agencies control)
4. Body Demanding Results	Shareholders, Owners	Not always clear. Donors? Clients?
5. Use of Monitoring Data	Seen as the key to higher efficiency	Sometimes in low demand for policy (“it pays to be ignorant”)

Source: Adapted from Gonzalo Hernandez (CONEVAL)

# In HD, spending is weakly correlated with outcomes...



- ...it's not trivial to establish and maintain good monitoring systems in the public sector
- ...but it's crucial for good impact evaluation

## And other key learning/accountability functions

- Tracking sector outcomes, MDGs etc (across time)
- Benchmarking system performance (across space)

# Objectives of this session

1. Why monitor?

2. How to monitor

➤ Using a RESULTS chain

3. How to measure what you want to know

Selecting indicators

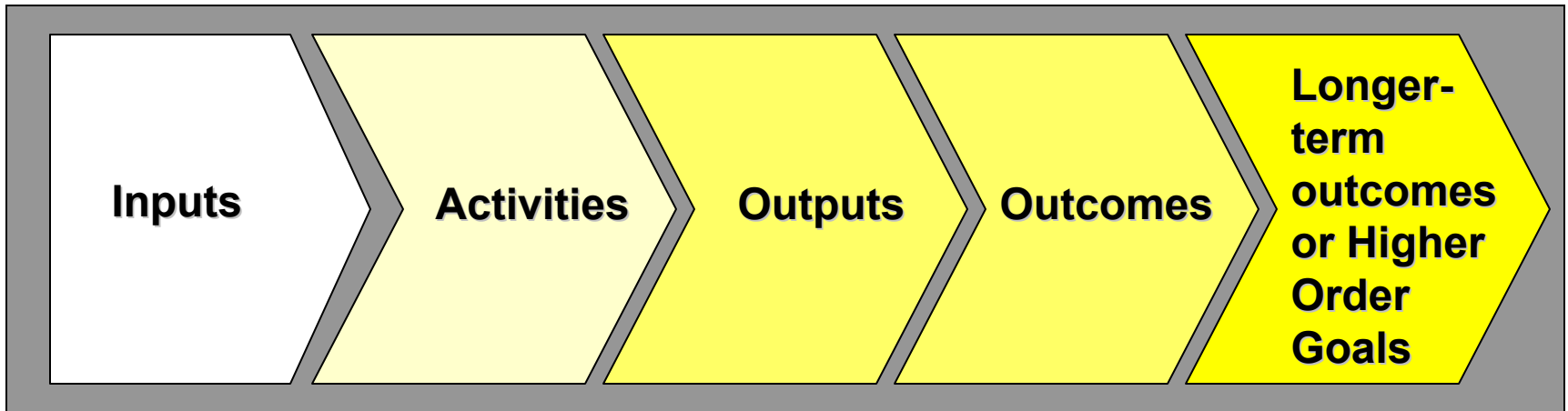
Setting up data collection

# Using a Results Chain

## A results chain answers 3 questions

- **What** are the intended results of the program?
- **How** will we achieve the intended results?
- **How** will we know we have achieved the intended results?

# The Results Chain in a Typical Program



Financial, human, and other resources mobilized to support activities

Actions taken or work performed to convert inputs into specific outputs

Project deliverables *within the control of implementing agency*  
**SUPPLY SIDE**

Use of outputs by beneficiaries and stakeholders *outside the control of implementing agency*  
**DEMAND SIDE**

Changes in outcomes that have multiple drivers

Budget  
Staffing

Training  
Studies  
Construction

Training plan completed  
Cash transfer delivered  
Road constructed  
School built

**New practices adopted**  
**Family nutrition improves**  
**Use of the road**  
**Children learn more**

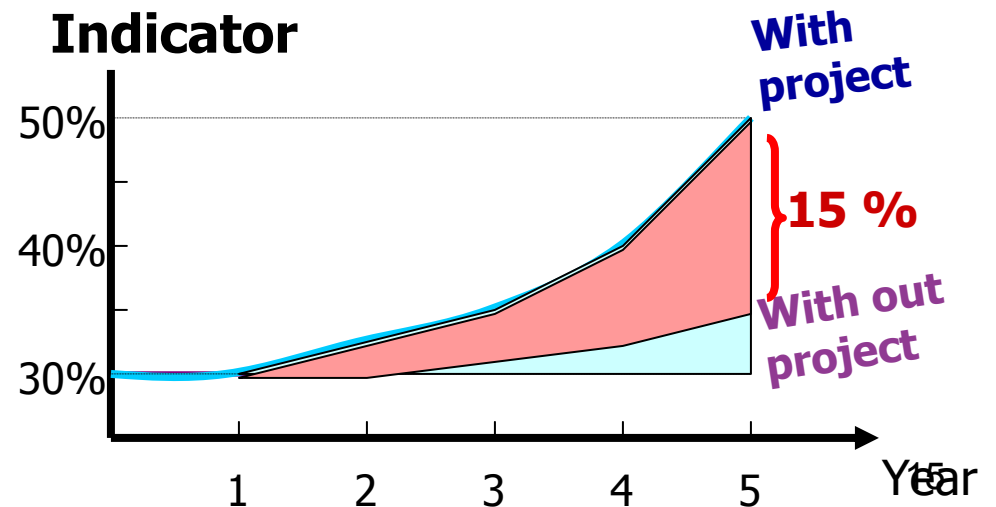
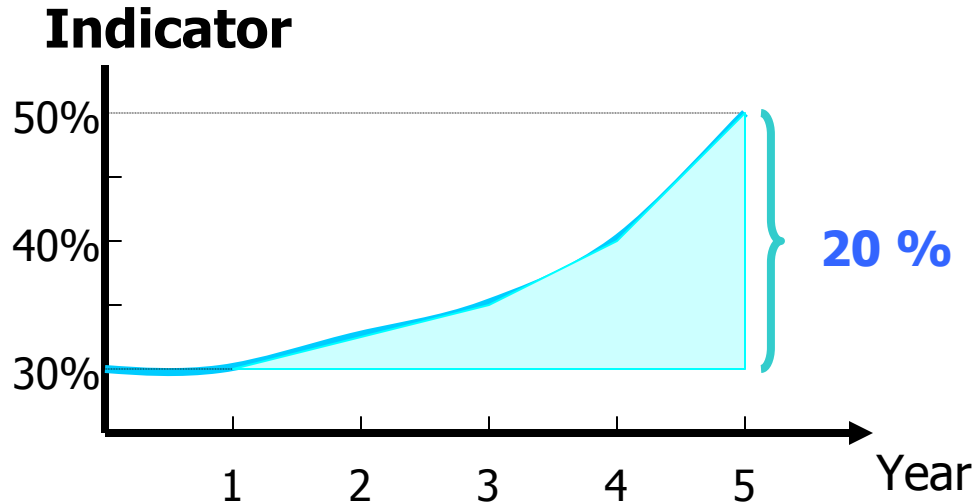
Poverty reduced  
Income inequality reduced  
Labor productivity increased

Implementation

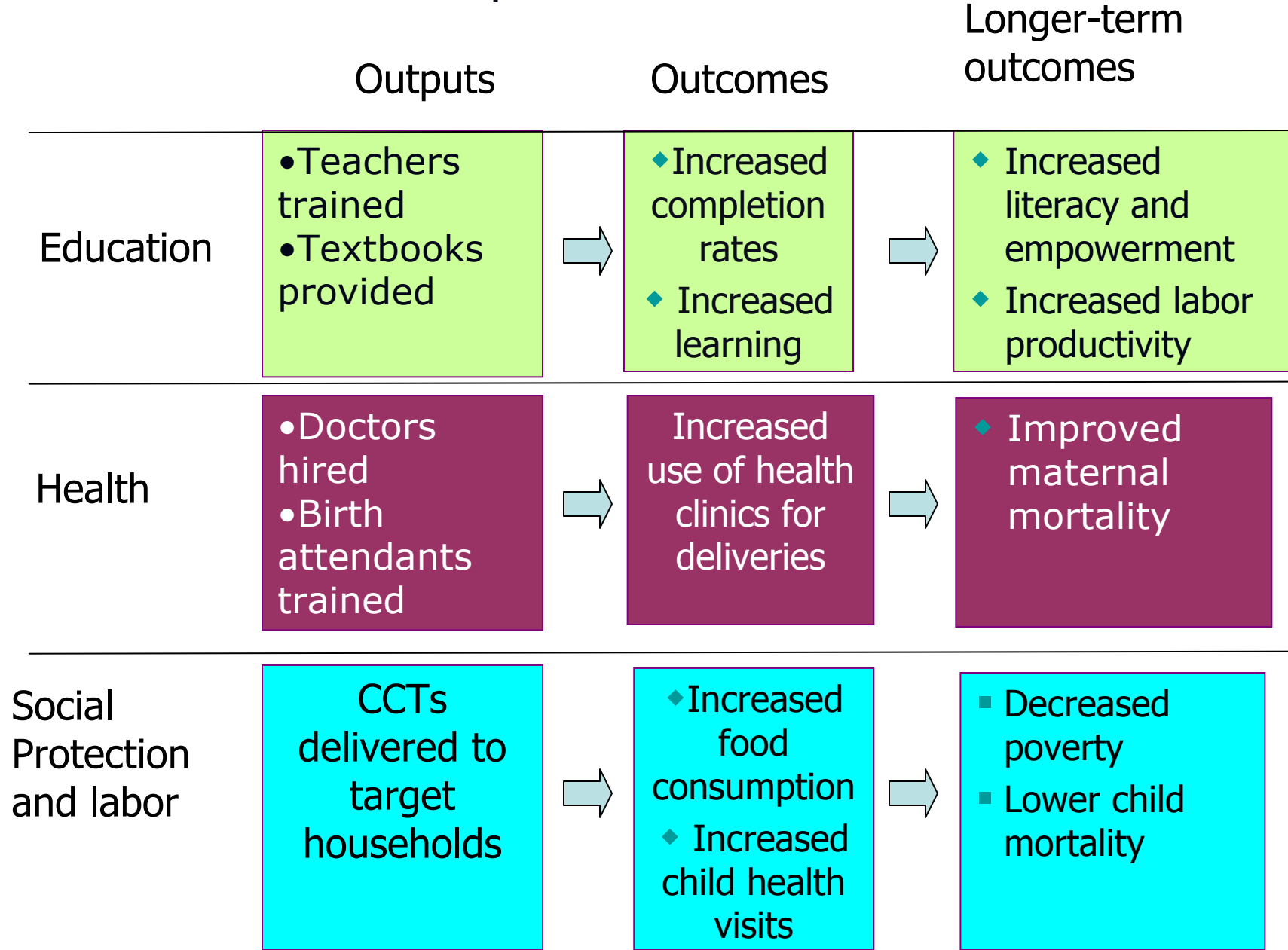
Results

Results-based Management

# Monitor the outcome indicator, evaluate the impact



# Examples of Results Chains



Identify the sequence of inputs, activities, outputs and outcomes

Example  
of Results  
Chain  
(Health  
Sector)

- 1. Information is available for parents about the importance of breast feeding**
- 2. Children in community healthier**
- 3. Fewer children are having diarrheal diseases**
- 4. Mothers breast feeding rather than using formula**
- 5. New funds available to implement a health project to reduce child mortality rates**
- 6. Design information campaigns on the importance of breast feeding**

Identify the sequence of inputs, activities, outputs and outcomes

Example  
of Results  
Chain  
(Health  
Sector)

- 5. New funds available to implement a health project to reduce child mortality rates - INPUT**
- 6. Design information campaigns on the importance of breast feeding – ACTIVITY**
- 1. Information is available for parents about the importance of breast feeding – OUTPUT**
- 4. Mothers breastfeeding rather than using formula – OUTCOME**
- 3. Fewer children are having diarrheal diseases – OUTCOME**
- 2. Children in community healthier – (higher level) OUTCOME**

# Objectives of this session

1. Why monitor?
2. How to monitor /RESULTS chain
3. How to measure what you want to know
  - Selecting indicators
  - Timing data collection
  - Mixed methods

# Operationalizing terms

General

**Construct -- Health**



**Indicator -- Body  
Temperature**

Specific

**Measure -- Degrees  
-- Celsius  
-- Fahrenheit**

# Selecting Indicators for an Education Project

Construct	Dimension	Outcomes	Indicators
School quality	Teacher capacity	Improvements in teacher practice	<ul style="list-style-type: none"> <li>• Number of teachers with certification</li> <li>• Child/parent satisfaction</li> </ul>
	Access to education	Improvements in access to school and classes	<ul style="list-style-type: none"> <li>• Enrollment ratio</li> <li>• Completion rate</li> <li>• Attendance ratio</li> <li>• Pupil/teacher ratio</li> </ul>
	Children's learning	Improvements in children's knowledge	Test achievement <ul style="list-style-type: none"> <li>• Increase in mean test scores</li> <li>• Reduced variance in test scores</li> </ul>

# Identifying good indicators: **SMART**

**S**pecific; **M**easurable; **A**ttributable; **R**ealistic; **T**argeted



✓ **Specific:** measure as closely as possible what you want to know

Outcome

Indicator

Parents insure that children treated for malaria

1. Increased utilization of clinics
  2. Increased use of malaria drugs
- Which indicator is more specific ?

✓ **Measurable:** be clear about how it will be measured

1. % of health centers without stocks of drugs x, y & z for more than a week at a time
2. % of health centers with availability of drugs

Which indicator is measurable ?

# SMART indicators

✓ **Attributable:** logically and closely linked to a program's efforts

1. Life expectancy
2. % of children fully immunized at 1 year

*Which indicator is attributable ?*

✓ **Realistic:** data obtainable at reasonable cost, frequency, and accuracy

1. HIV prevalence among 15-24 year old pregnant women
2. HIV prevalence among the total population

*Which indicator is more realistic?*

✓ **Targeted:** Specific to the program's target group

1. Percent increase in employment
2. Percent increase in employment of graduates of technical training center X in the first year after completion of training<sup>23</sup>

*Which indicator is targeted ?*

# Criteria for selecting indicators

- Collectable
- ***Relevant***
- ***Valid***
- ***Reliable***
- Reasonable cost/burden
- Privacy/confidentiality
- Timely
- Comprehensive

# Validity

- Make sure indicators is an accurate representation of the reality you want to capture
- If it's a proxy indicator, is it genuinely linked to what you want to measure?
  - Eg, teacher certification not good proxy for teacher quality

# Reliability

- Are you measuring the same thing over time?
- Is the indicator “stable”?
  - eg, measuring costs over time (nominal vs. real) prices
  - Student learning (is the assessment standardized?)

# Relevance

- Are you measuring what really counts in the transformation of outputs to outcomes?
- Use indicators that are meaningful
- Avoid trap of measuring what is easy to measure

# An example from Brazil: teacher bonus pay reform

Intervention: 3 months' salary bonus for all staff of schools which meet their annual performance targets – based on student flows and student test scores in grades 4 and 8

What do we want to know?

- Does the bonus reform improve school results over time – student learning and student flows?
- How does its impact per unit of dollar spent compare with other alternative education policies/programs?

# An example from Brazil: teacher bonus pay reform

What else do we want to know?

- Does the bonus motivate teachers to perform better?
  - Reduce absenteeism?
  - Improve teaching practice? Time on task?
- Does the bonus motivate schools to collaborate and plan better?
  - Increase time spent sharing practice and developing ideas to improve school results?
- Does the bonus stimulate adverse behaviors (teaching to the test, reduced time for non-tested subjects, blame and free-riding)?

Are the teachers reacting like this?

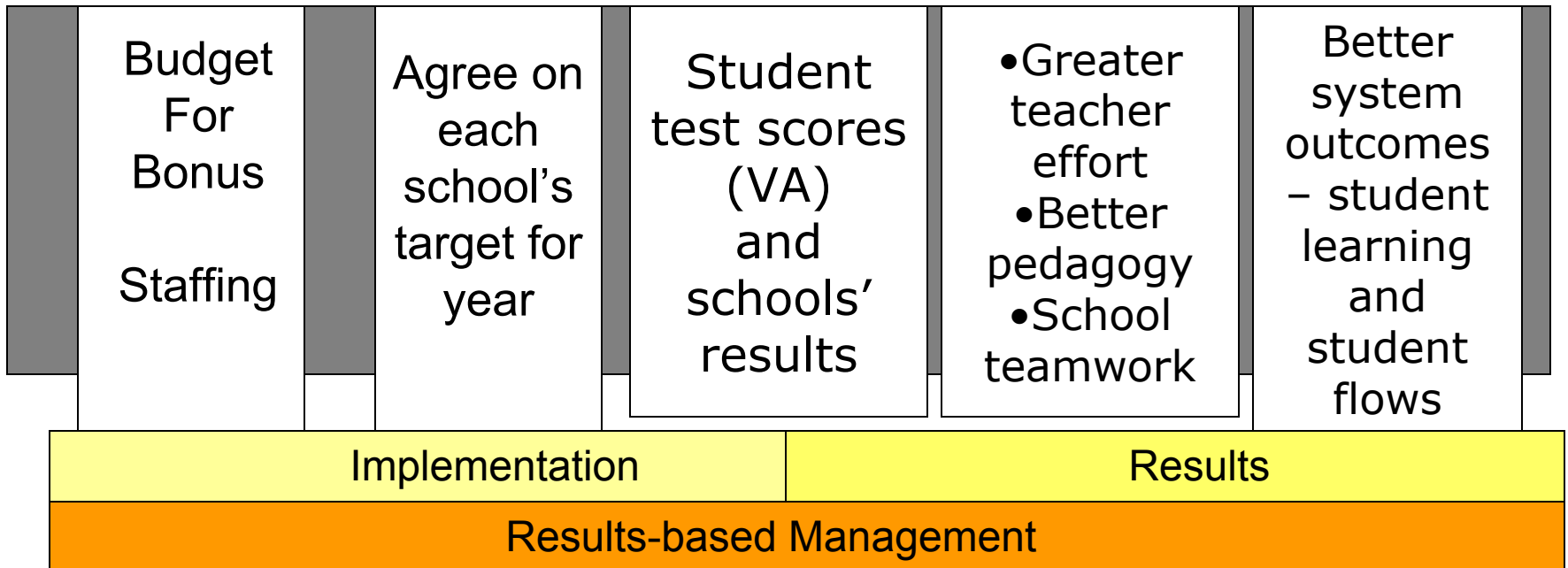
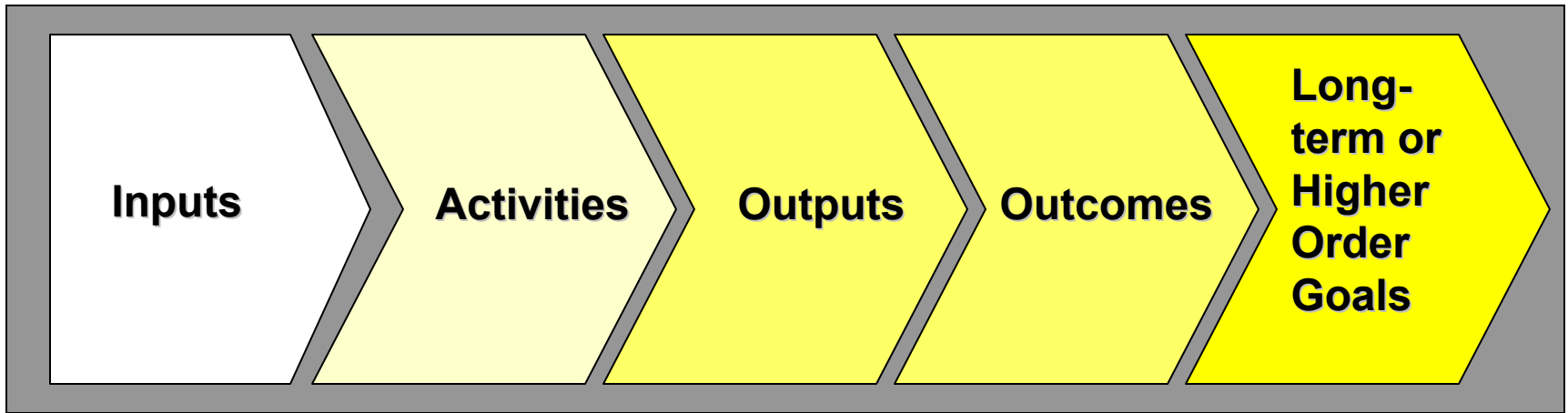


.....or this?



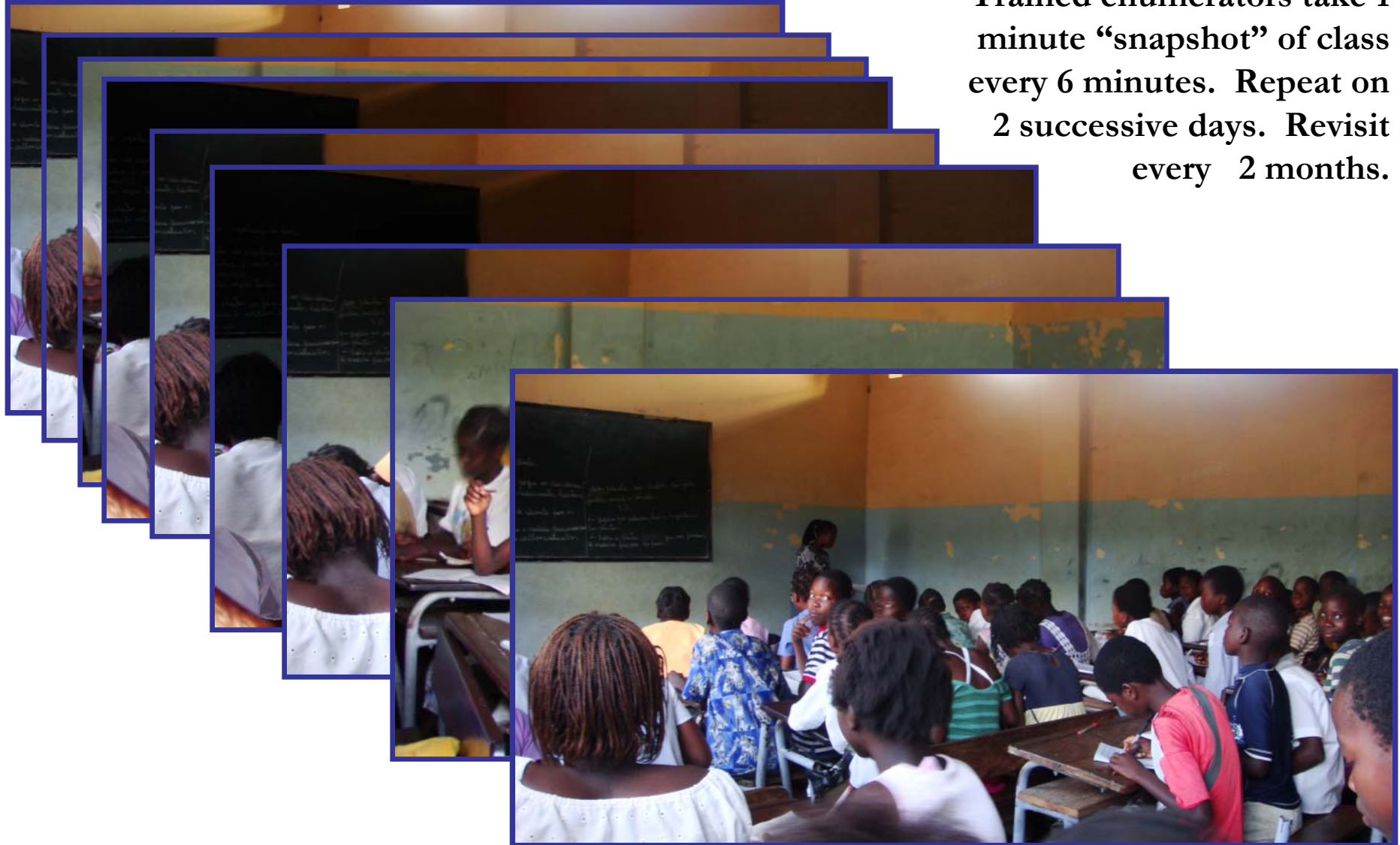
05.01.2006 11:43

# Brazil Teacher Incentives Reform



# Stallings Classroom Snapshot

Trained enumerators take 1 minute “snapshot” of class every 6 minutes. Repeat on 2 successive days. Revisit every 2 months.



# Excerpt from Classroom Snapshot

CLASSROOM OBSERVATION SNAPSHOT								
ACTIVITY		MATERIAL						
		NO MATERIAL	TEXTBOOK	NOTEBOOK / WRITING MATERIAL	BLACK BOARD	VISUAL AIDES	MANIPULATIVES	COOPERATIVE
1. READING ALOUD	T	1 S L E	1 S L E	1 S L E	1 S L E	1 S L E	1 S L E	S L E
	I	1 S L	1 S L	1 S L	1 S L	1 S L	1 S L	S L
	CHECK IF CHORAL READING <input type="checkbox"/>							

T line: Indicates activities that involve the teacher.

I line: Indicates activities that involve the student & not the teacher.

1, S, L, E: Indicate one individual, a small, large group and entire class respectively.

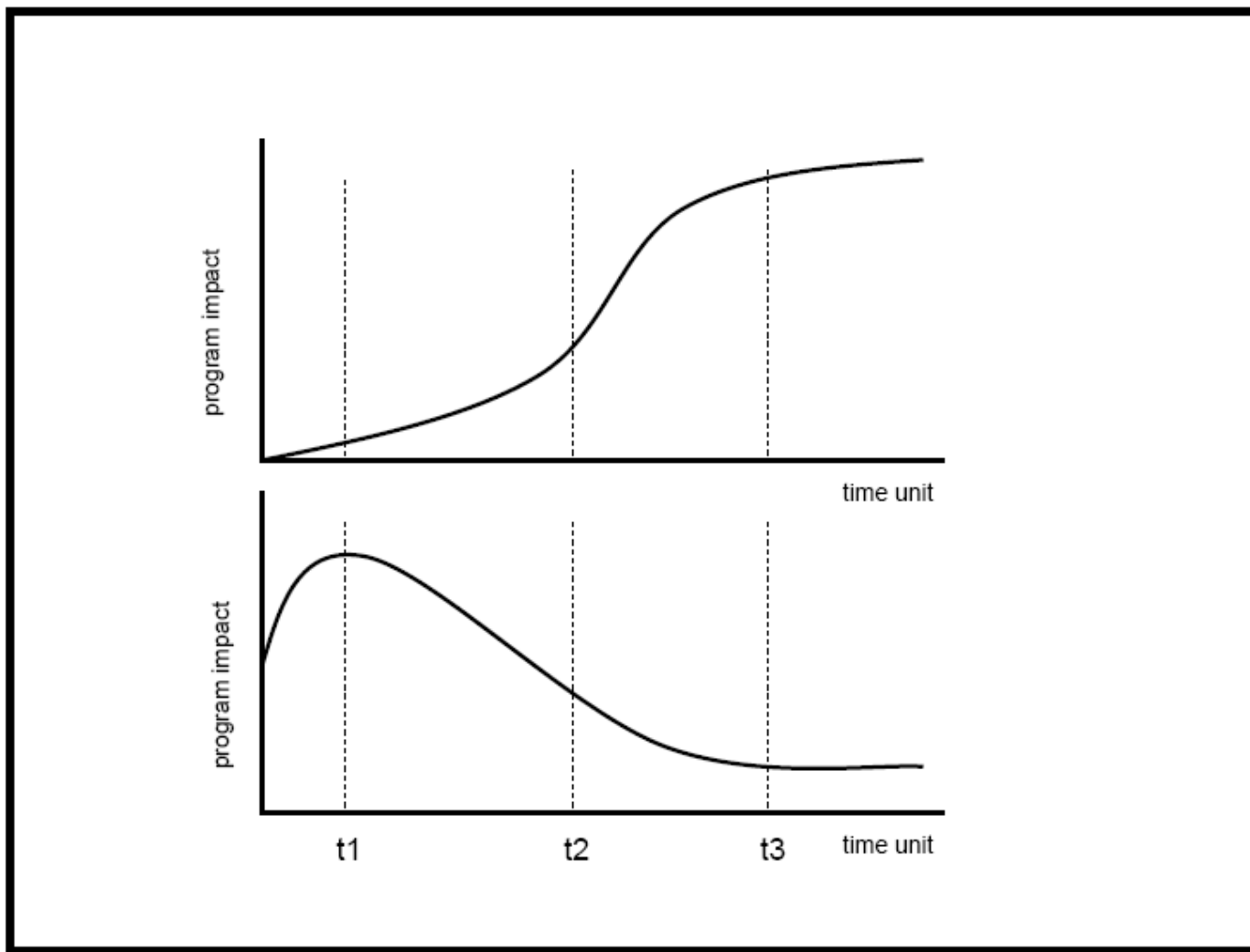
# Coding Sheet

CLASSROOM OBSERVATION SNAPSHOT								
ACTIVITY	MATERIAL							
		NO MATERIAL	TEXTBOOK	NOTEBOOK / WRITING MATERIAL	BLACK BOARD	VISUAL AIDES	MANIPULATIVES	COOPERATIVE
1. READING ALOUD	T	1 S L E	1 S L E	1 S L E	1 S L E	1 S L E	1 S L E	S L E
	I	1 S L	1 S L	1 S L	1 S L	1 S L	1 S L	S L
	CHECK IF CHORAL READING <input type="checkbox"/>							
2. LECTURE	T	1 S L E	1 S L E	1 S L E	1 S L E	1 S L E	1 S L E	S L E
	I	1 S L	1 S L	1 S L	1 S L	1 S L	1 S L	S L
3. DISCUSSION	T	1 S L E	1 S L E	1 S L E	1 S L E	1 S L E	1 S L E	S L E
	I	1 S L	1 S L	1 S L	1 S L	1 S L	1 S L	S L
4. PRACTICE & DRILL	T	1 S L E	1 S L E	1 S L E	1 S L E	1 S L E	1 S L E	S L E
	I	1 S L	1 S L	1 S L	1 S L	1 S L	1 S L	S L
5. FEEDBACK	T	1 S L E	1 S L E	1 S L E	1 S L E	1 S L E	1 S L E	S L E
	I	1 S L	1 S L	1 S L	1 S L	1 S L	1 S L	S L
6. ASSIGNMENT	T	1 S L E	1 S L E	1 S L E	1 S L E	1 S L E	1 S L E	S L E
	I	1 S L	1 S L	1 S L	1 S L	1 S L	1 S L	S L
7. COPYING	T	1 S L E	1 S L E	1 S L E	1 S L E	1 S L E	1 S L E	S L E
	I	1 S L	1 S L	1 S L	1 S L	1 S L	1 S L	S L
8. VERBAL INSTRUCTION	T	1 S L E	1 S L E	1 S L E	1 S L E	1 S L E	1 S L E	S L E
	I	1 S L	1 S L	1 S L	1 S L	1 S L	1 S L	S L
9. SOCIAL INTERACTION	T	1 S L E						
	I	1 S L						
10. STUDENT(S) UNINVOLVED	I	1 S L						
11. DISCIPLINE	T	1 S L E						
12. CLASSROOM MANAGEMENT	T	1 S L E						
	I	1 S L						
13. CLASSROOM MANAGEMENT ALONE/TRANSITION ACTIVITIES					T			
14. TEACHER SOCIAL INTERACTION OR TEACHER UNINVOLVED					T			
15. TEACHER OUT OF THE ROOM					T			

# Sample Results

Percentage of Class Time Used . . .					
	<i>Teacher-Related</i>				<i>Student-Related</i>
Region/ Country	Discipline	Interactive Learning	Passive Learning	Room Cleaning	Students Not Engaged
Brazil (Pernambuco)	1.8%	18.2%	52.9%	18.6%	28.5%
Ghana	1.4%	25.9%	52.5%	12.5%	35.0%
Morocco	4.2%	9.4%	62.9%	20.0%	17.8%
Tunisia	0.9%	9.6%	61.7%	26.3%	12.0%

Figure 1. The timing of evaluations can affect impact estimates



Source: **King and Behrman, "Timing and Duration of Exposure in Evaluations of Social Programs", World Bank WPS 4686, August 2008**

# Timing

- Implementation lags
  - Start evaluation when “complete” intervention delivered
- Effects can deepen with time
  - Learning by providers and beneficiaries
  - Outcomes can be cumulative/deepen
  - Unforeseen outcomes can arise
- Effects can dissipate with time
  - control areas catch up
  - effects wash out (India computers in classroom)
- Understand the time path of program
  - plan data collection accordingly

# Mixed method designs

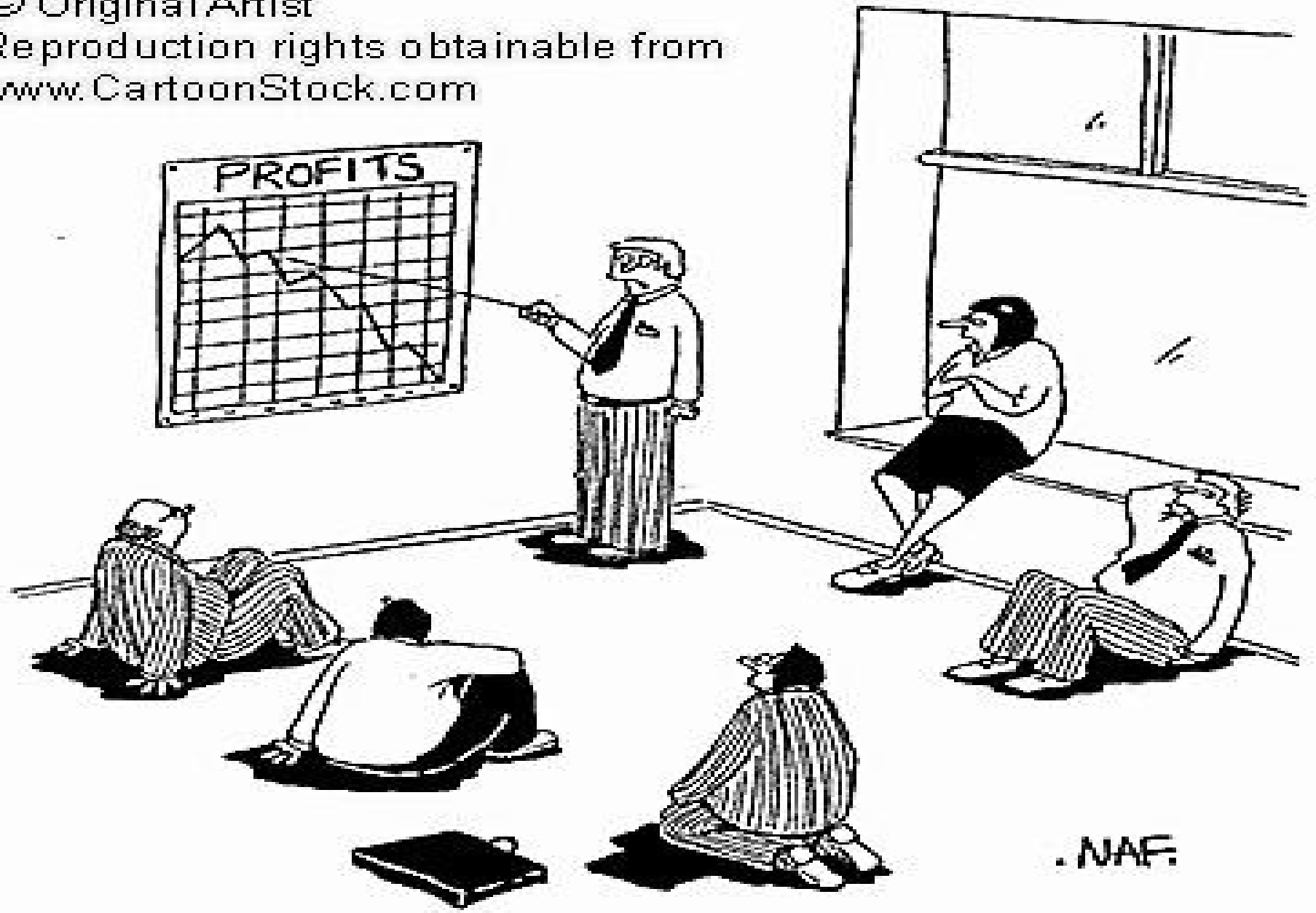
- Combining quantitative and qualitative methods at all stages of the data collection makes for stronger M & E
- Qualitative methods
  - Focus groups
  - Case studies of outliers
- Combines depth (qualitative) and breadth and statistical certainty (quantitative)
- Strengthens validity through triangulation

# THE END

You cannot have a first rate impact evaluation without a good monitoring system

....you need to understanding how the program was implemented!

© Original Artist  
Reproduction rights obtainable from  
[www.CartoonStock.com](http://www.CartoonStock.com)



"And this is where we had to start selling the furniture."

©Cartoonbank.com



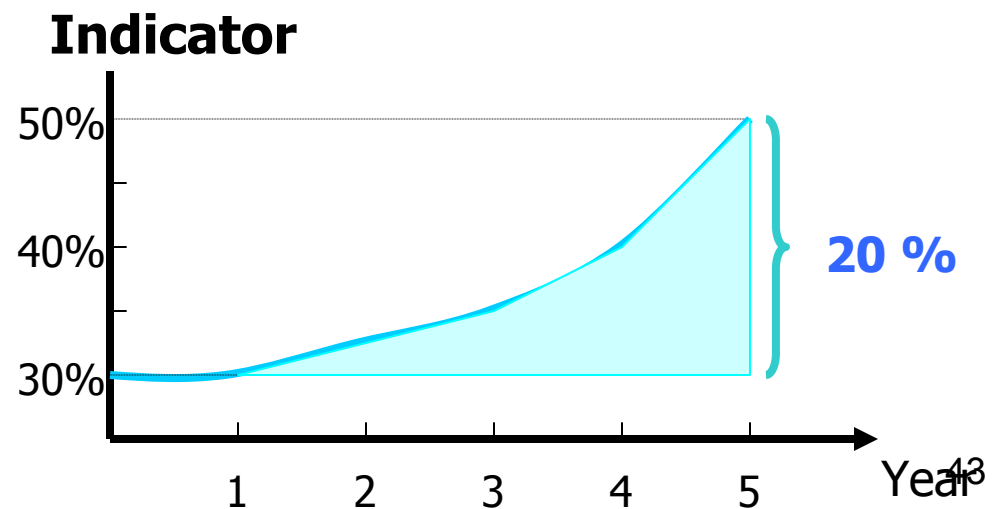
*"My question is: Are we making an impact?"*

# Monitoring



**Ongoing, real time data collection on inputs, outputs and outcomes to track implementation**

- Tracks changes in indicators over time**
- Tracks actuals compared to targets**
- Establishes whether target beneficiaries are being reached**
- Identifies implementation bottlenecks and problems for management attention**



# Evaluation



Ex post analysis to explain how outcomes were or were not achieved

- Establishes causality
- Validates hypotheses
- Explores unintended outcomes
- Provides accurate measure of program outcomes
- Provides evidence base for better policy

