

Theory of Change and Measurement

Ilf Bencheikh
Deputy Executive Director
J-PAL Europe



Lecture Overview

- Theory of change
- Sources of measurement
- Measurement concepts

Lecture Overview

- **Theory of change**
- Sources of measurement
- Measurement concepts

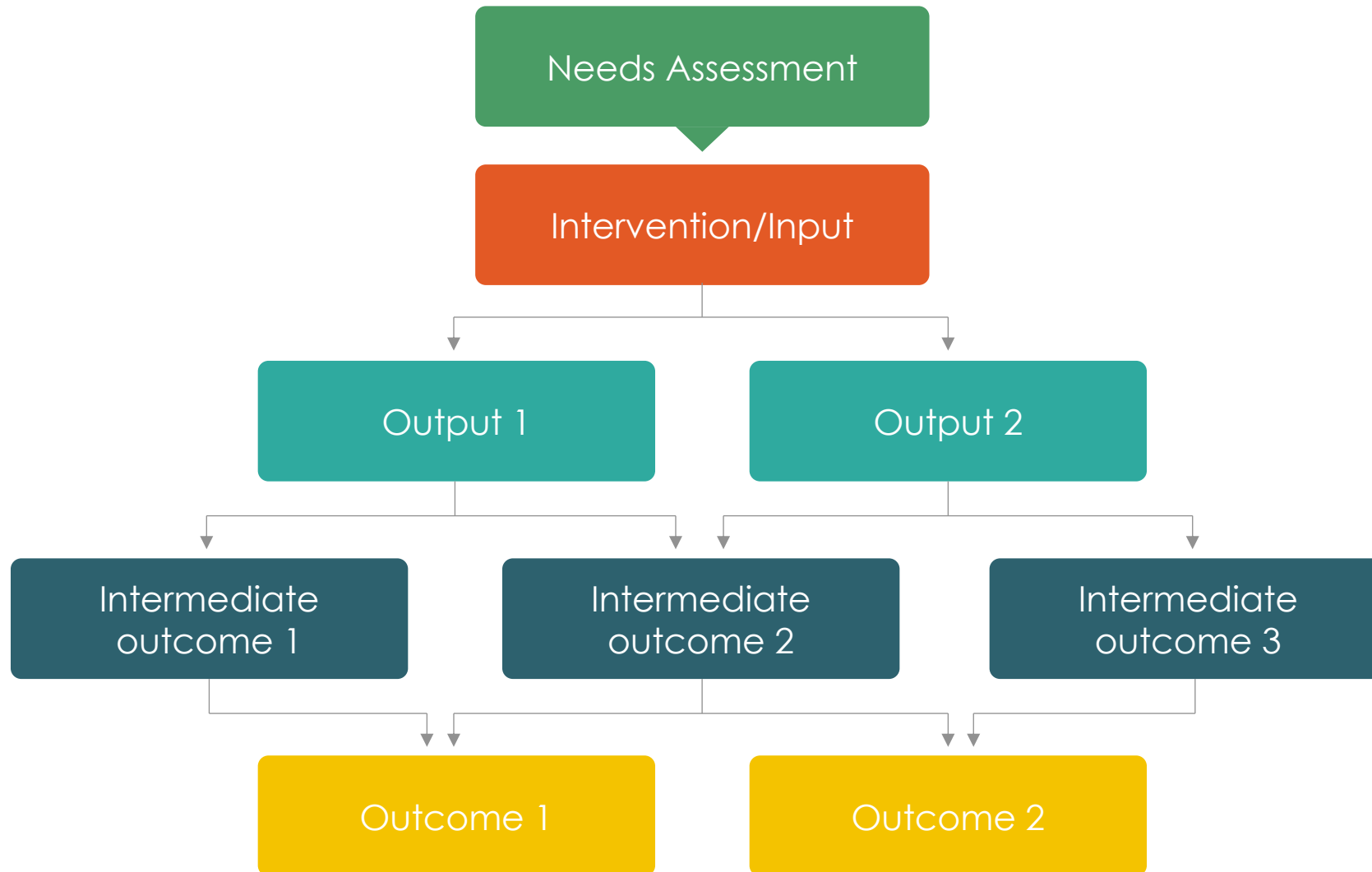
Components of Program Evaluation

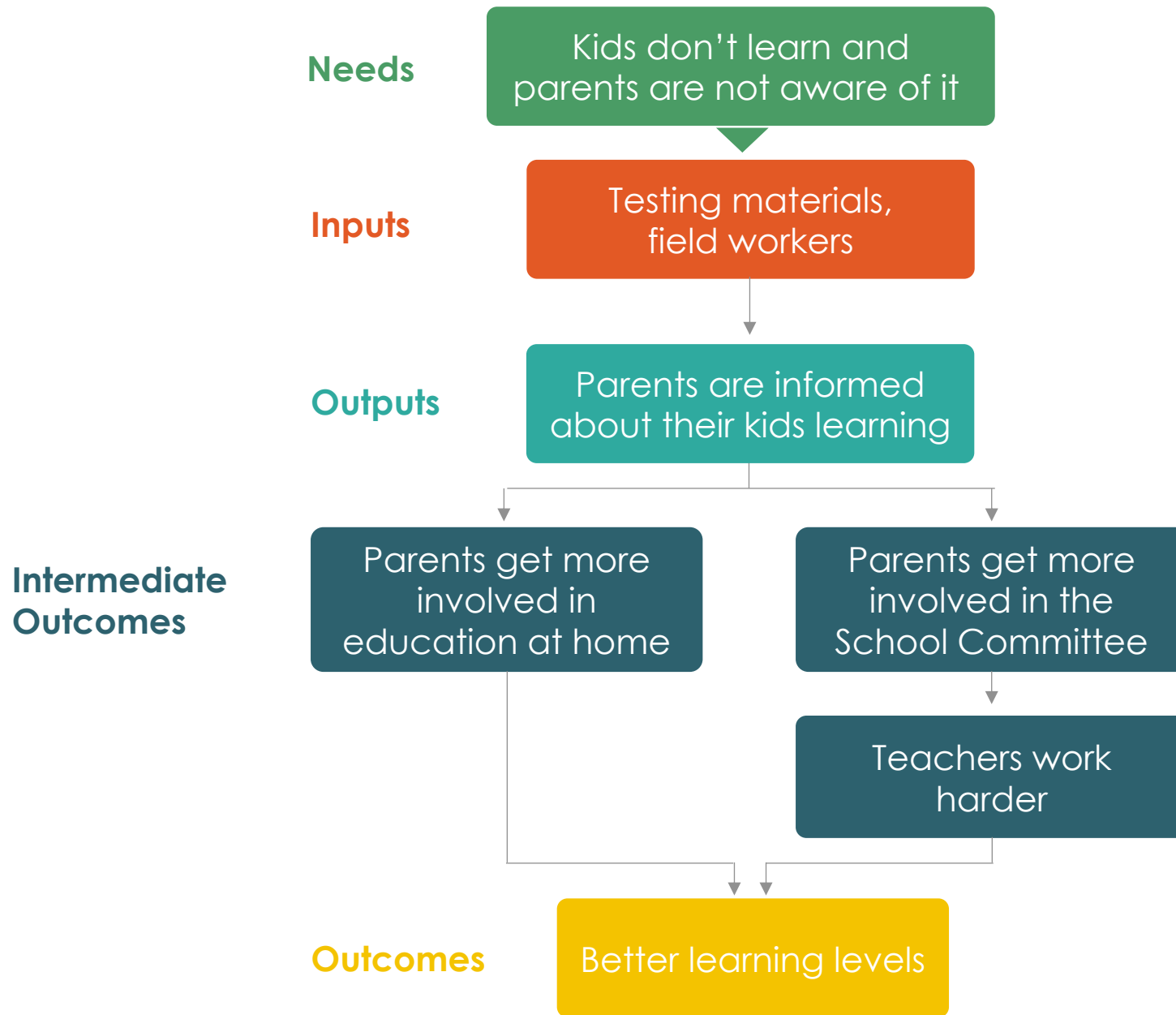


Definition

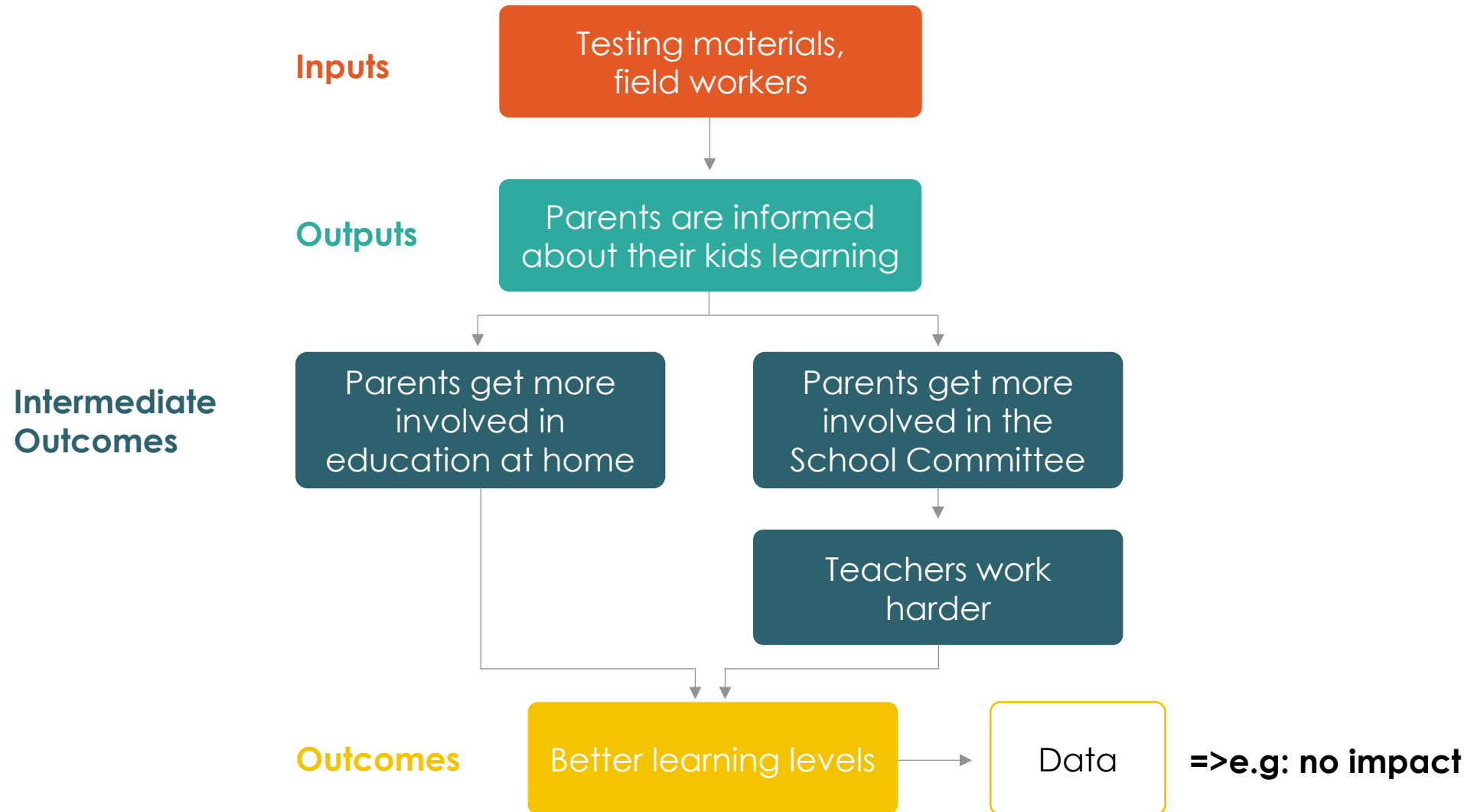
- The **Theory of Change** defines a set of connected building blocks, generally called inputs, outputs, and outcomes.
- It is depicted on a map that is **a graphic representation of the change process.**

Theory of Change

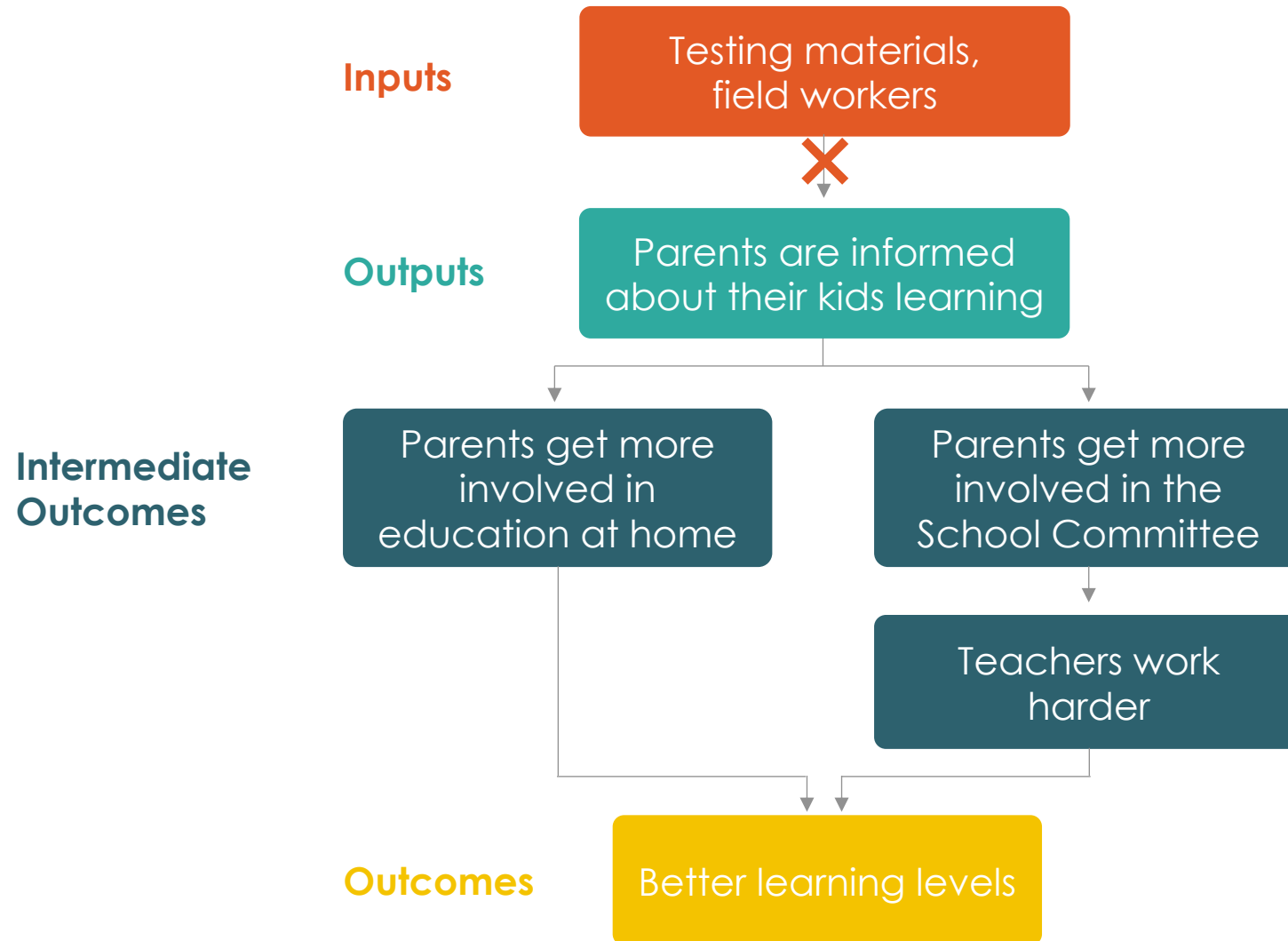




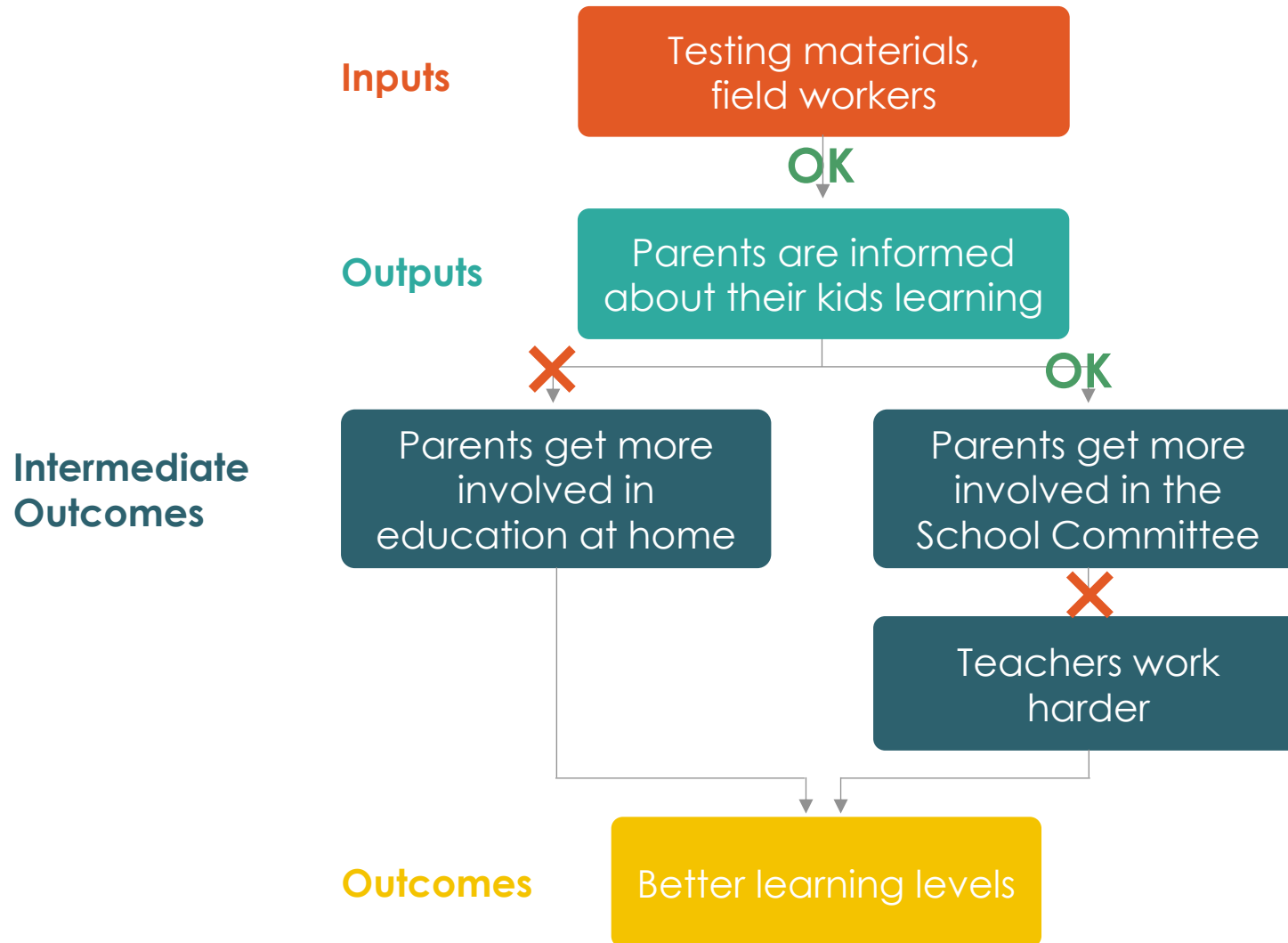
No final impact: what can we conclude?



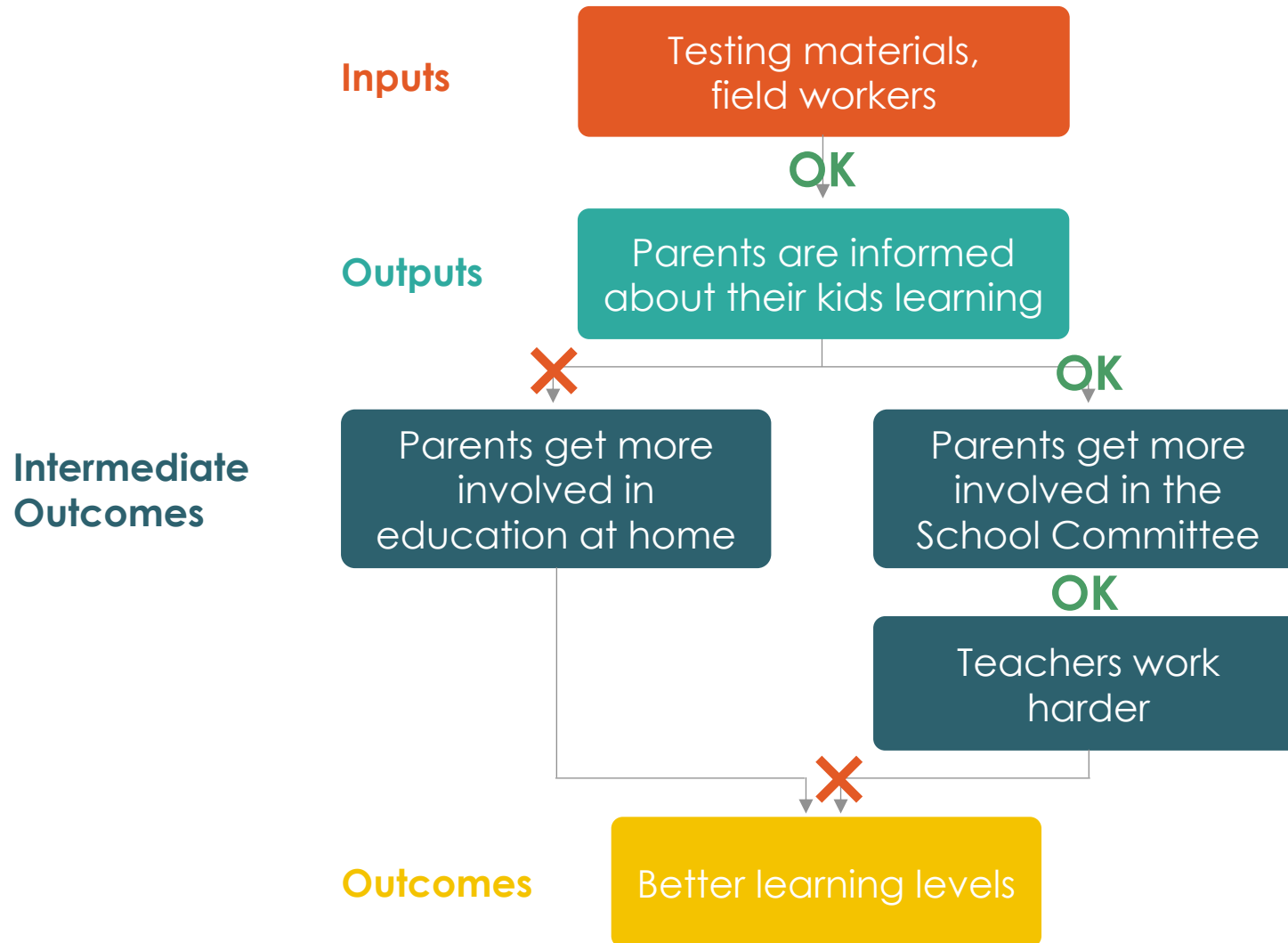
Scenario 1



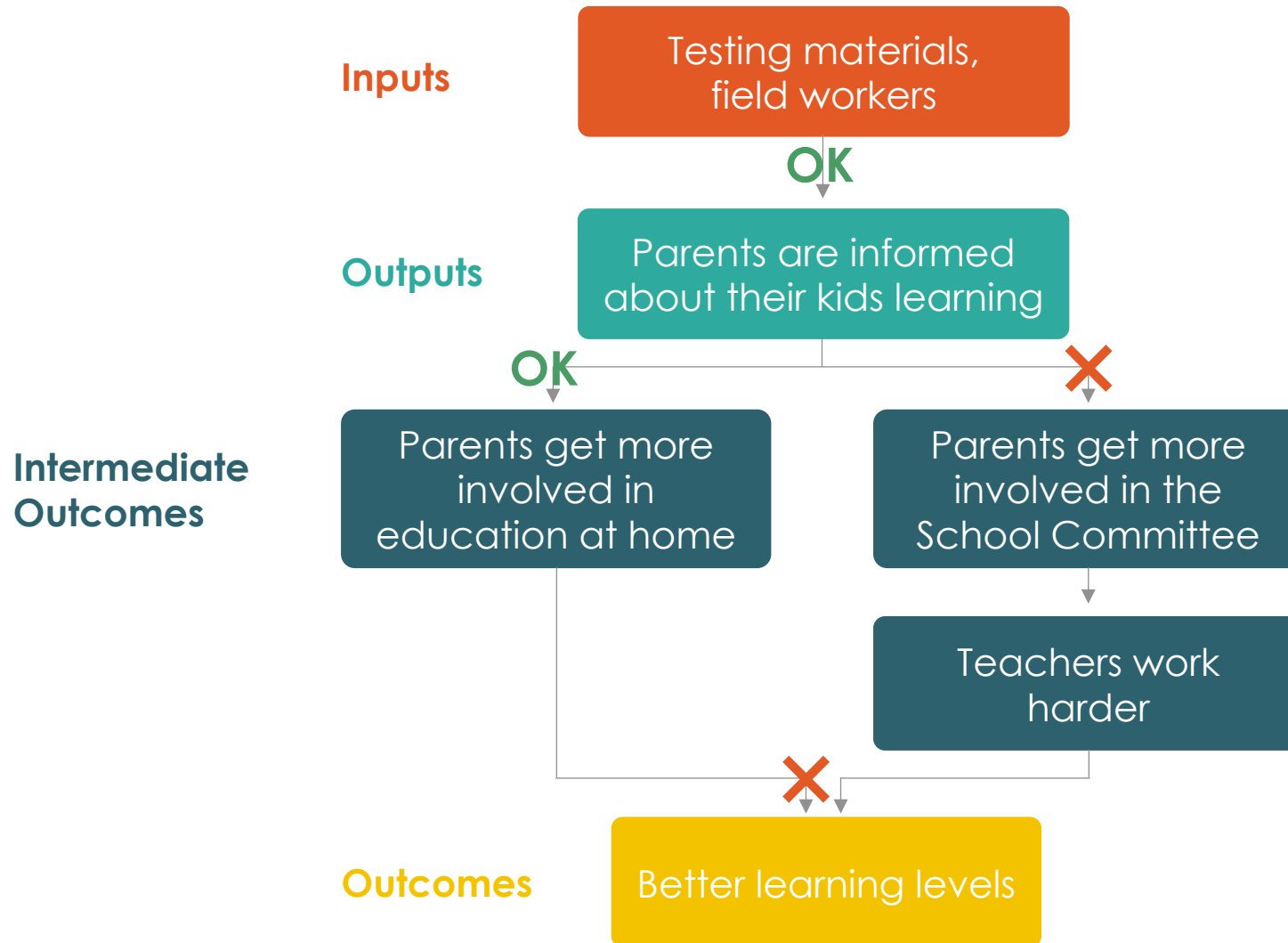
Scenario 2



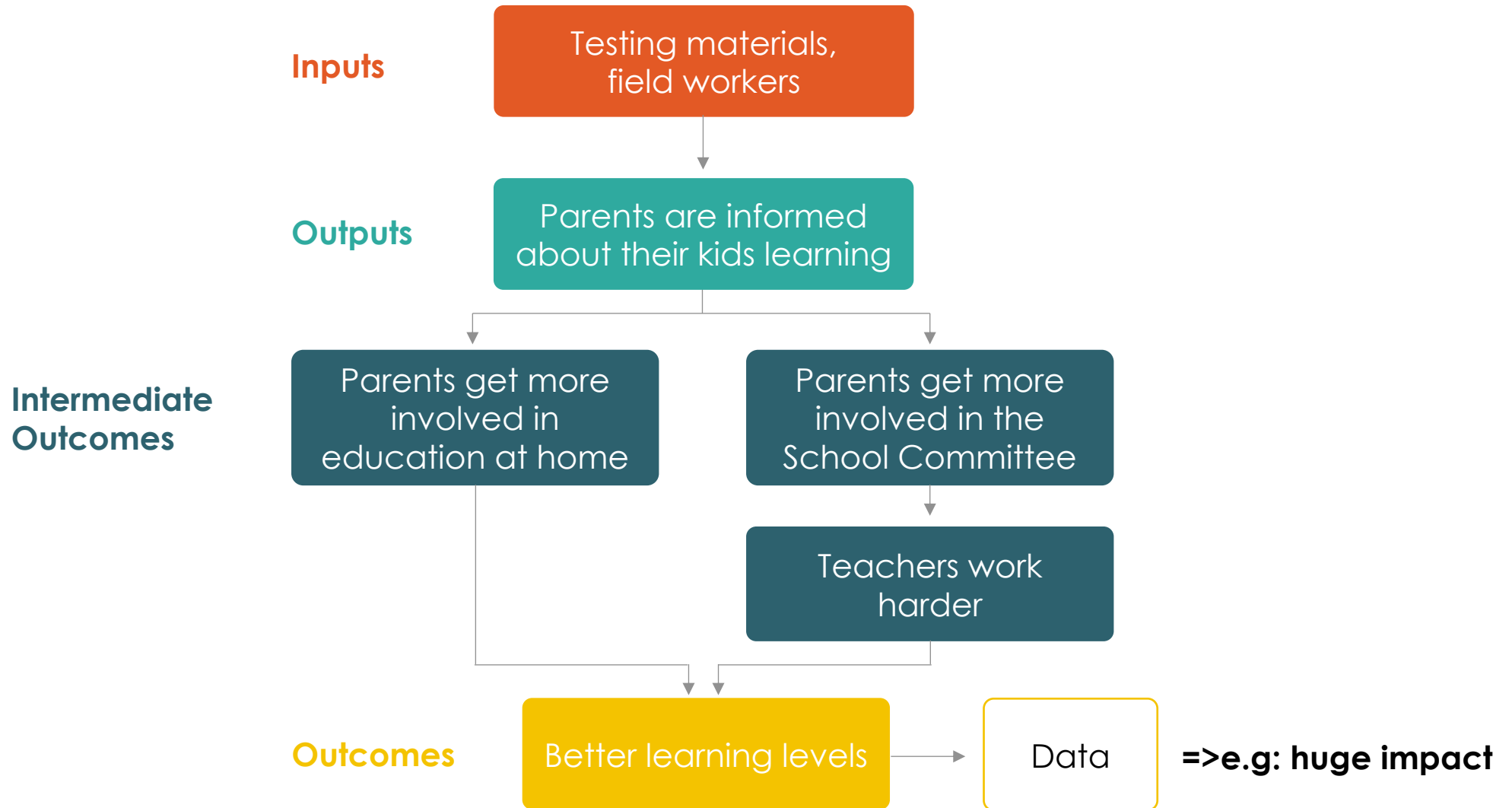
Scenario 3



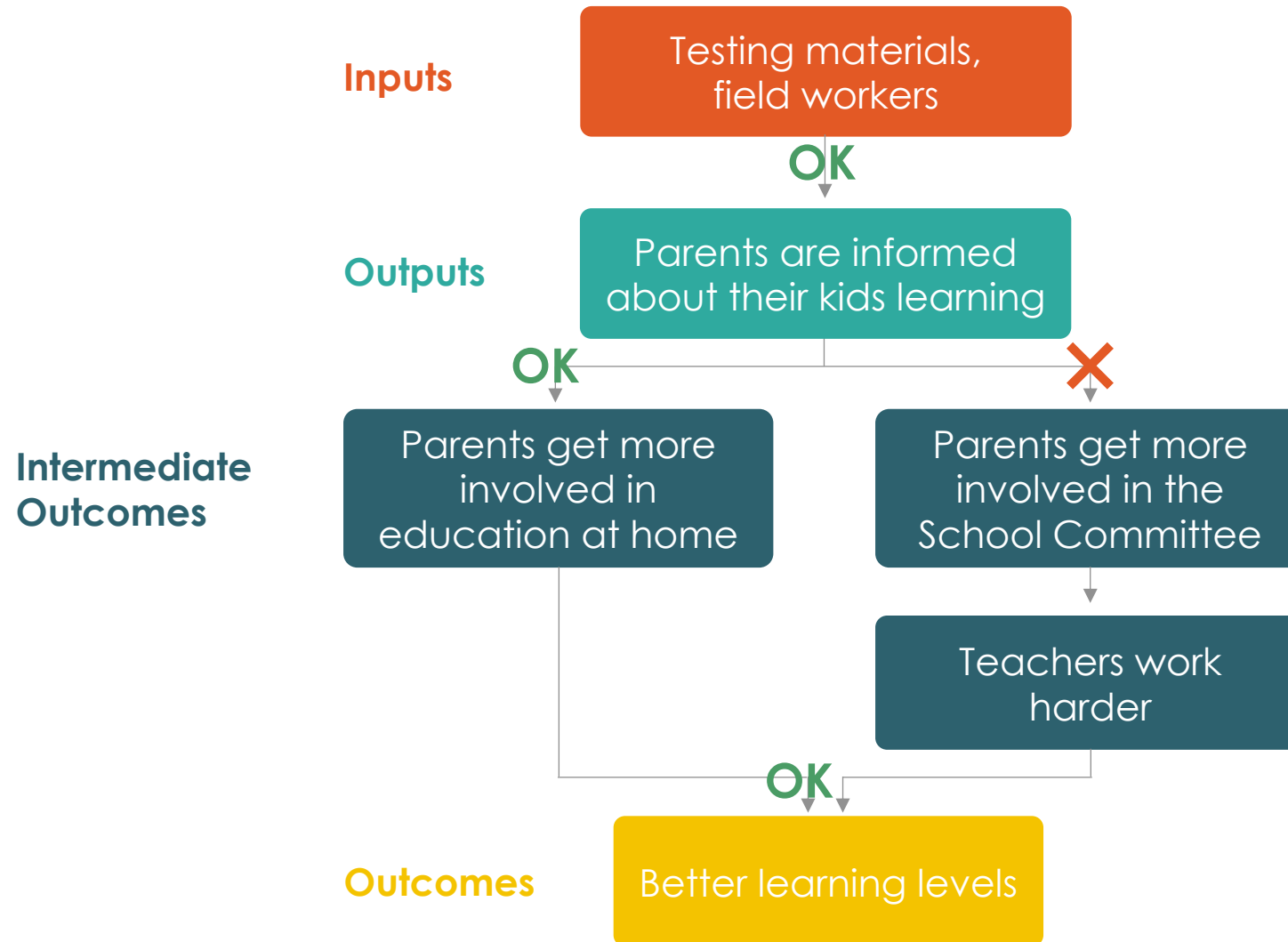
Scenario 4



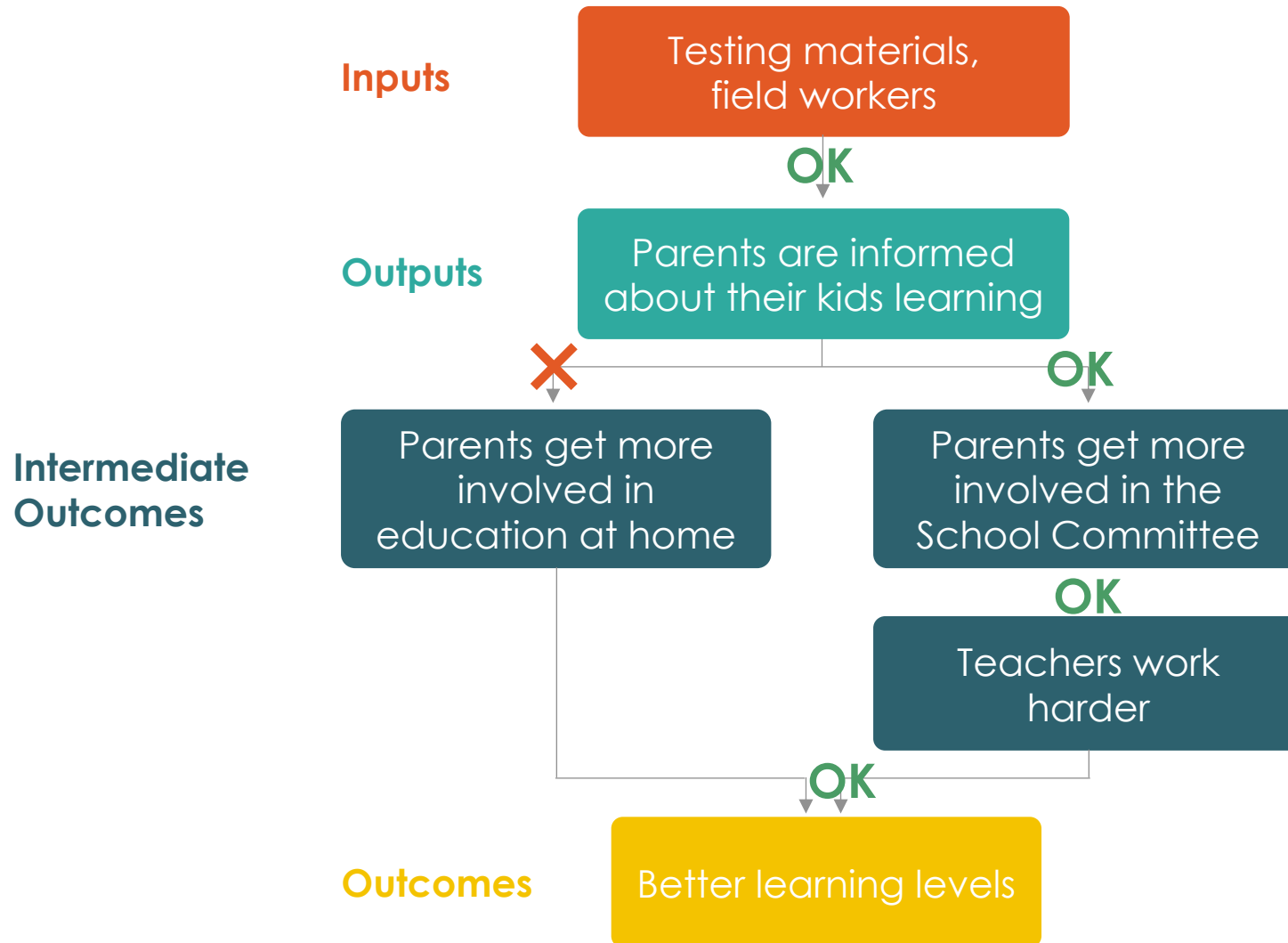
Huge impact



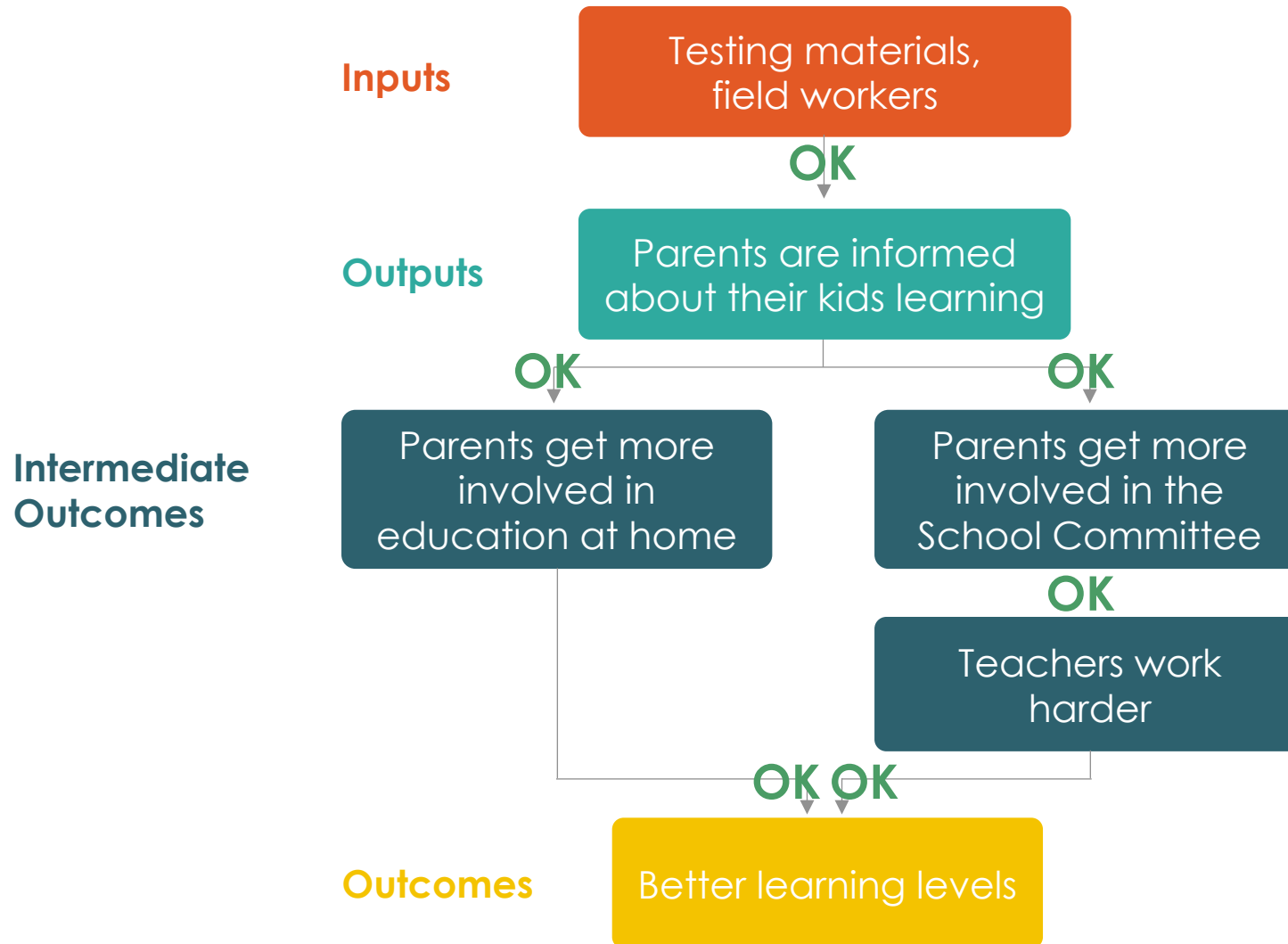
Scenario 1



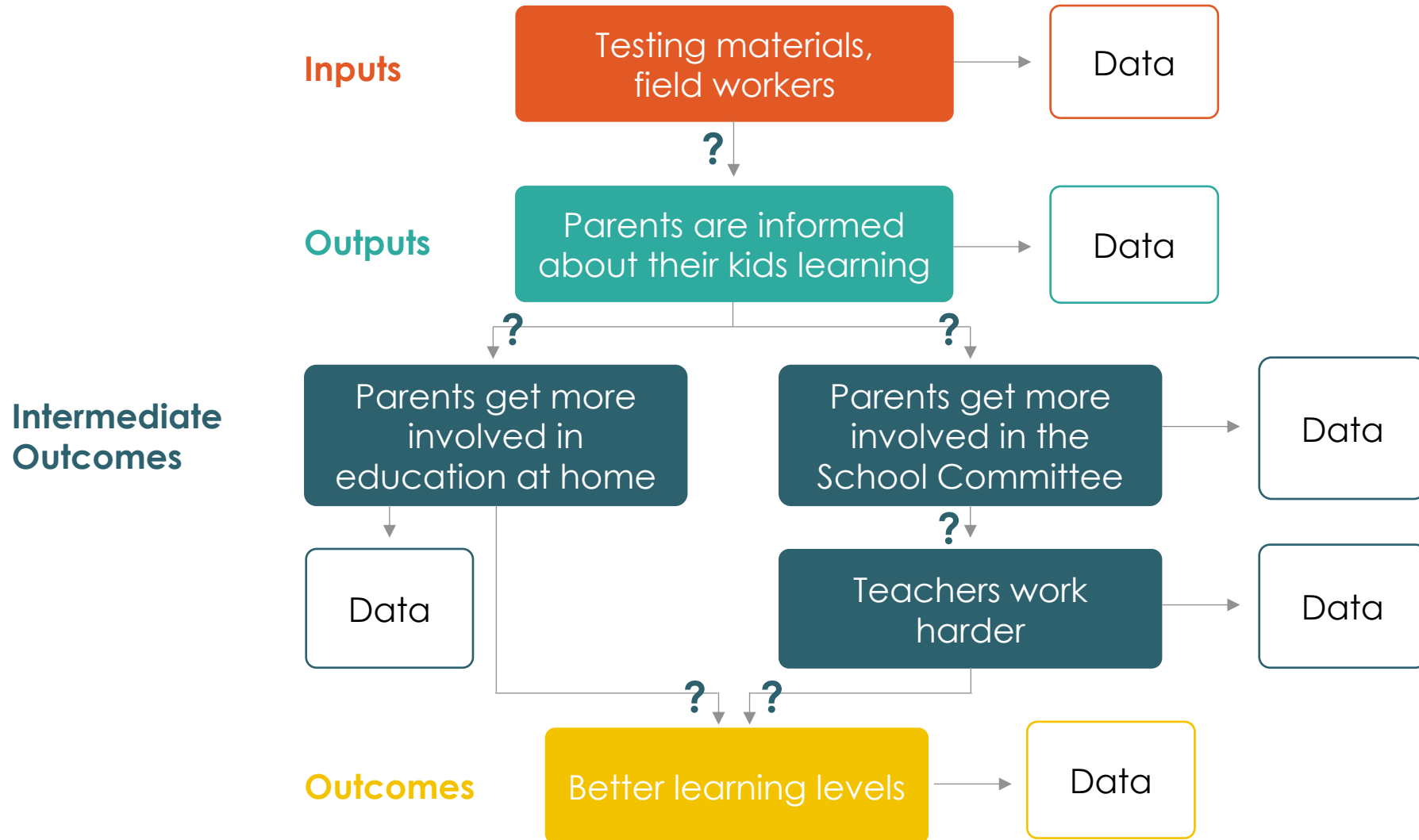
Scenario 2



Scenario 3



You need data at each stage



Why spend time on the theory of change?

1. Helps design the intervention

- Often done backwards
- Is each arrow really credible?

2. Helps design the evaluation by:

- Generating research questions
- Deciding which data to collect

By measuring the right intermediate variables, we can get “into the black box”

- Allows to understand the “why”, thus giving richer policy lessons
- Gives more generalizable knowledge

Lecture Overview

- Theory of change
- **Sources of measurement**
- Measurement concepts

Where can we get data?

- From existing sources (Secondary data)
 - Publicly available
 - Administrative data
 - Other secondary data
- Collected by researchers (Primary data)
 - Surveys
 - Non-survey methods



Photos: https://commons.wikimedia.org/wiki/File:Cuyahoga_County_US_Census_Form-Herbert_Birch_Kingston_1920.jpg;
[https://commons.wikimedia.org/wiki/File:US_Navy_090123-N-97607-004_Hospital_Corpsman_2nd_Class_Jennifer_Ross_files_medical_records_ aboard_the_aircraft_carrier_USS_Nimitz_\(CVN_68\).jpg](https://commons.wikimedia.org/wiki/File:US_Navy_090123-N-97607-004_Hospital_Corpsman_2nd_Class_Jennifer_Ross_files_medical_records_ aboard_the_aircraft_carrier_USS_Nimitz_(CVN_68).jpg)

Types and Sources of Data

Information reported by a person

Information about a person/household

- Cognition, anthropometrics
- Demographics
- Behavior, beliefs
- Patience, risk aversion, psychometrics
- Knowledge
- Income, expenditure

NOT about a person/household

- Farming inputs and outputs
- Quality of medical care
- Business income taxes

Automatically generated

- Bank transactions
- Phone data
- Sales records
- School/university records, criminal record

- Prices
- Weather, air quality
- Stock markets
- VAT records

Primary Data Collection

- Surveys
- Exams, tests, etc.
- Games
- Vignettes
- Direct observation
- Diaries/logs
- Focus groups
- Interviews



Primary Data: Modes

- Interviewer administered
 - Paper-based
 - Computer-assisted/ digital
 - Telephone-based
- Self-administered
 - Paper
 - Computer/digital



Administrative data

Information collected, used, and stored primarily for administrative (i.e., operational) purposes, rather than research purposes.

- Medical records
- Grade books
- Arrest records
- Bank account data
- Personnel records
- Log books



Why are administrative data useful?

The outcomes and metrics required for a study may already be tracked by a government or organization

- Available retrospectively
- Enable long-term follow-up
- May include near census of relevant population
- Reduce logistical burden and burden on subjects
- Often less expensive than surveys
- May reduce bias and error

Lecture Overview

- Theory of change
- Sources of measurement
- **Measurement concepts**

Concepts of measurement

Construct

The main concept being investigated. A construct is often abstract.

(E.g., learning, teacher effort, parental involvement)



Indicators

How you actually measure or “operationalize” your construct.

(e.g. test scores, teacher preparedness, hours dedicated to helping kids do homework)



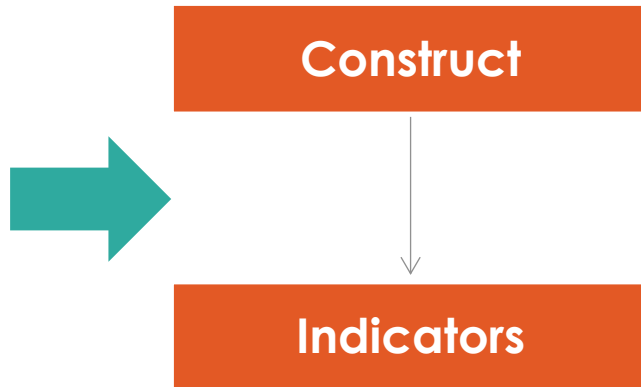
Data Collection



Data

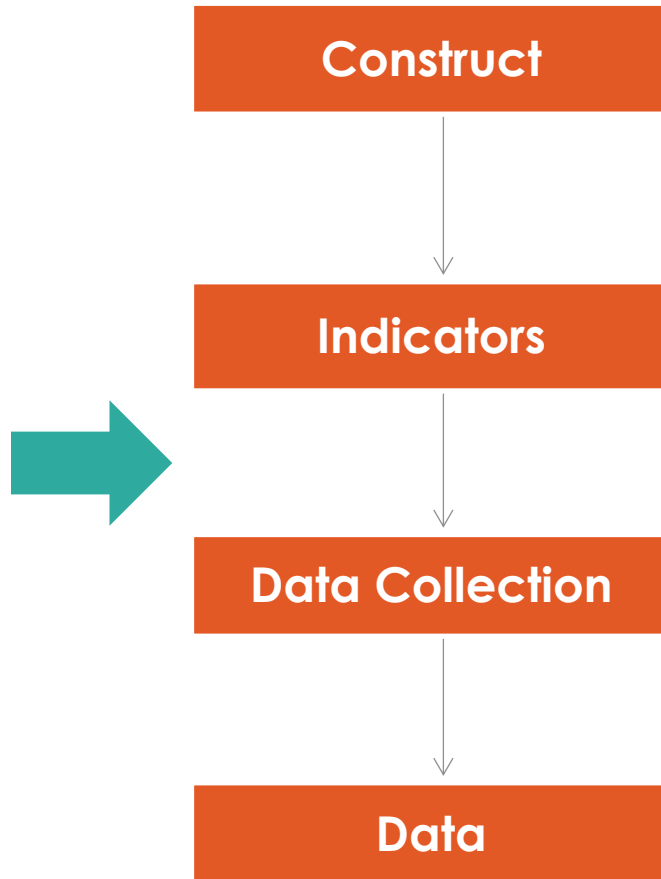
What we use to measure our indicators.

Validity



In theory: how well does the indicator map to the construct we are trying to measure?

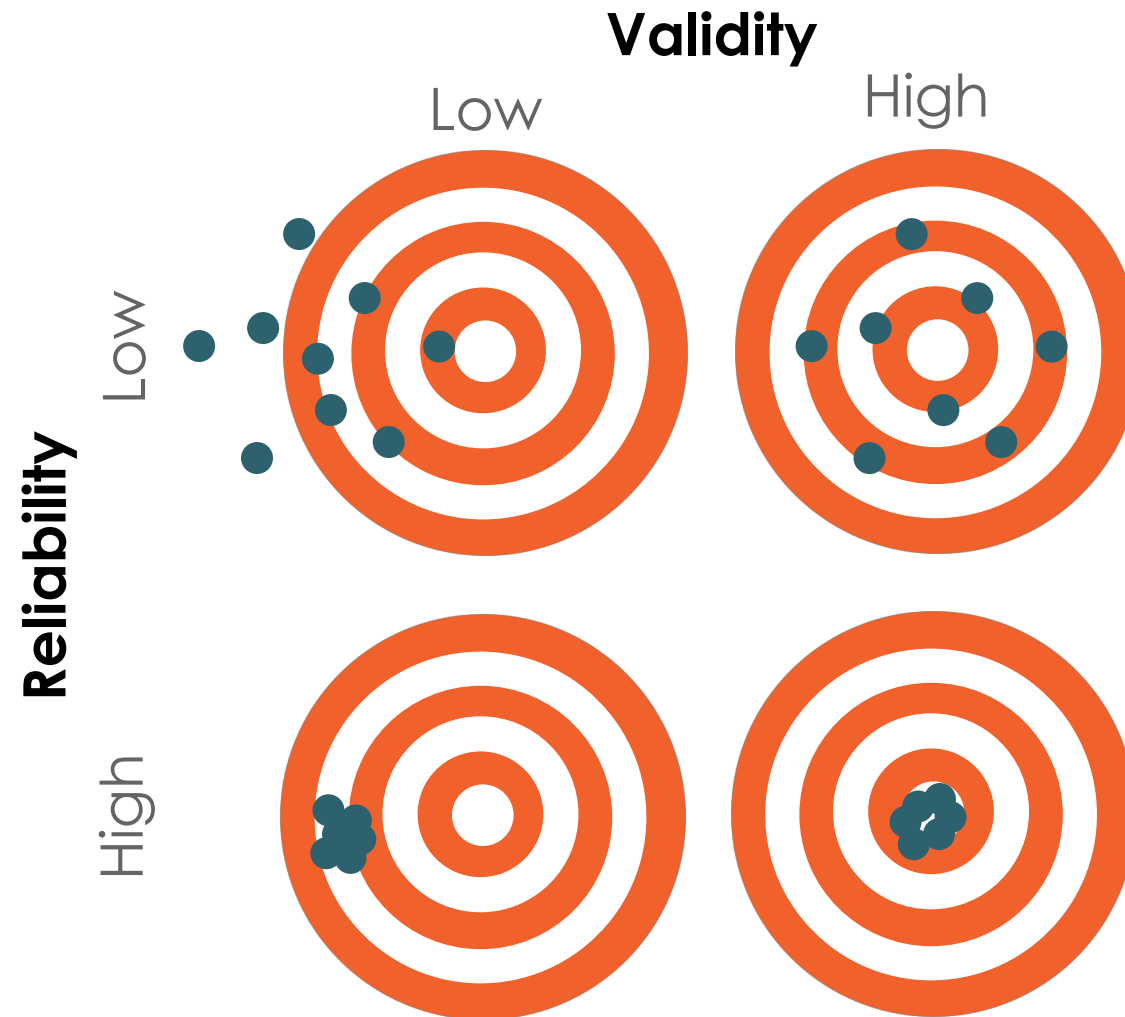
Reliability



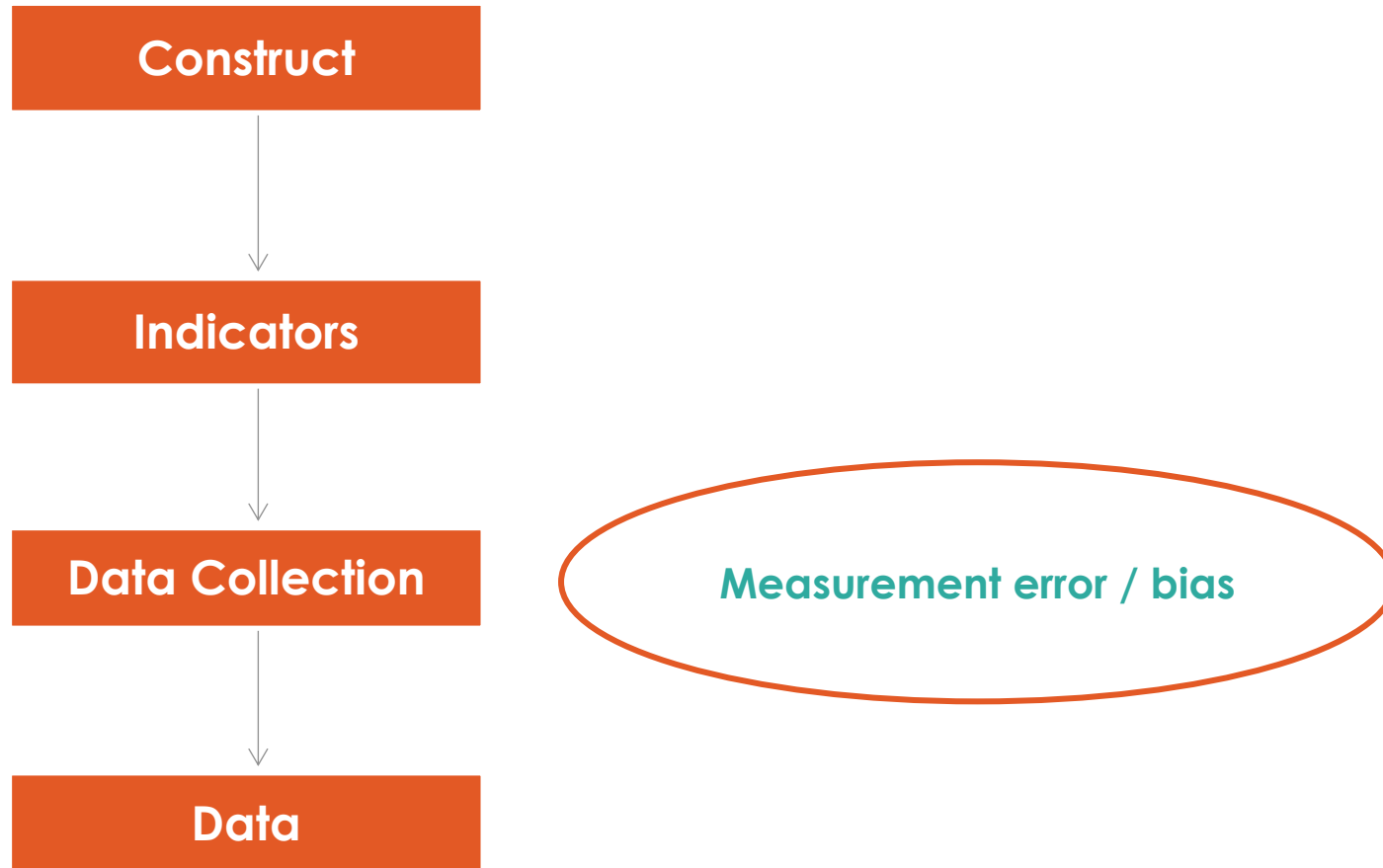
Is the indicator measured in a way that is consistent and precise?

Would we get the same data if we measured several times?

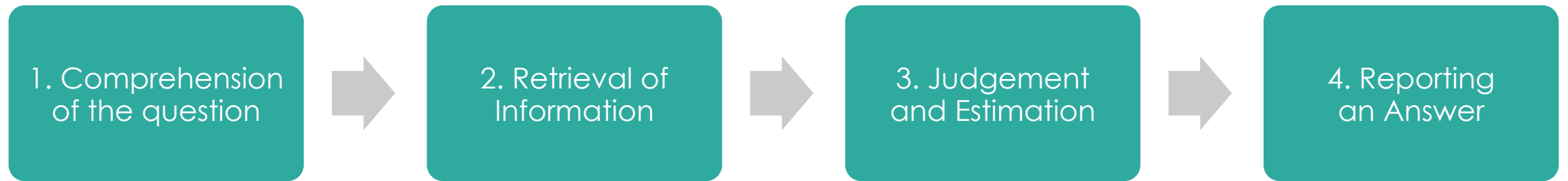
The Goals of Measurement



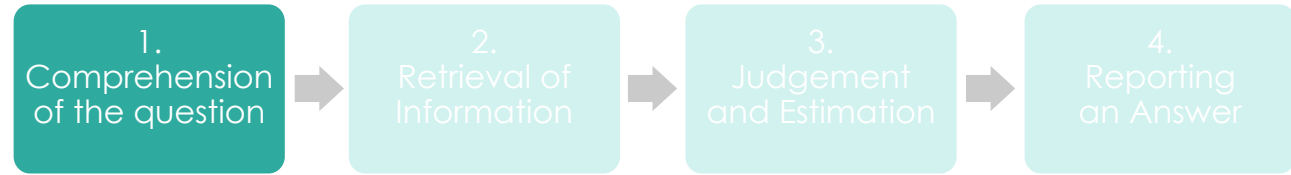
Data collection can go wrong



The Response Process



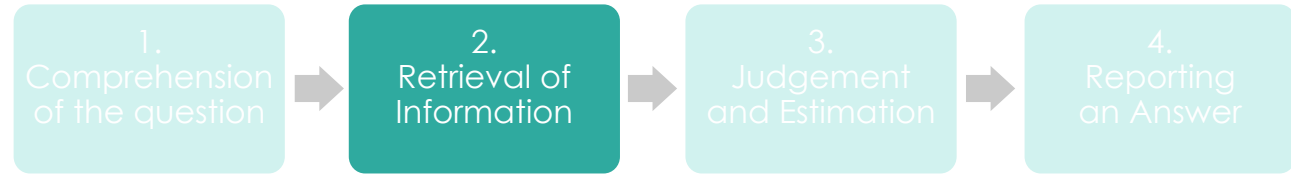
Step 1: Comprehension




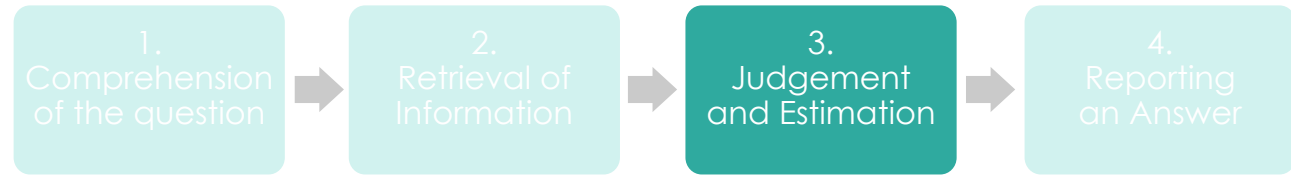
1.1 Total monthly income, before taxes



Step 2: Retrieval

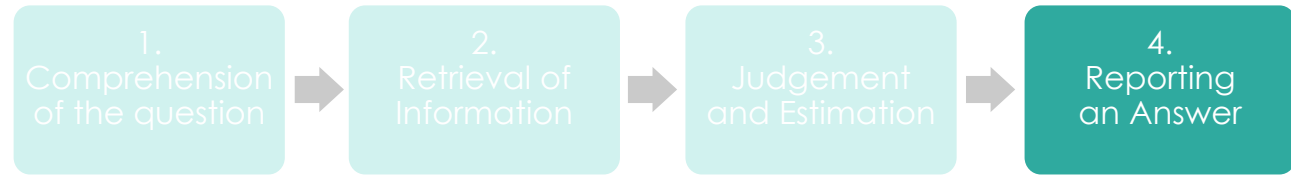


Step 3: Estimation/ Judgement



Social = \$ 200 per month
Workers' Compensation = 0
Pension = \$220 per month
What else??

Step 4: Response



1.1 Total monthly income, before taxes



Conclusion on data quality

- Particularly difficult to ensure the quality of survey data
=> Lot of energy and resources dedicated to that
- Good to combine survey data with other kind of data (direct observation, tests)...
- But those too can be imprecise / biased
- Admin data potentially subject to similar imprecision / biases as primary data

Time to illustrate with a case study!
Thank you!

