

Experimental Methods



THE WORLD BANK
IBRD • IDA



Strategic Impact
Evaluation Fund

Outline

1. What exactly do we mean by impact?
2. What can be learned from randomized experiments?



What do we mean by impact?

“An impact evaluation assesses changes in the well-being of individuals, households, communities or firms that can be attributed to a particular project, program or policy.”

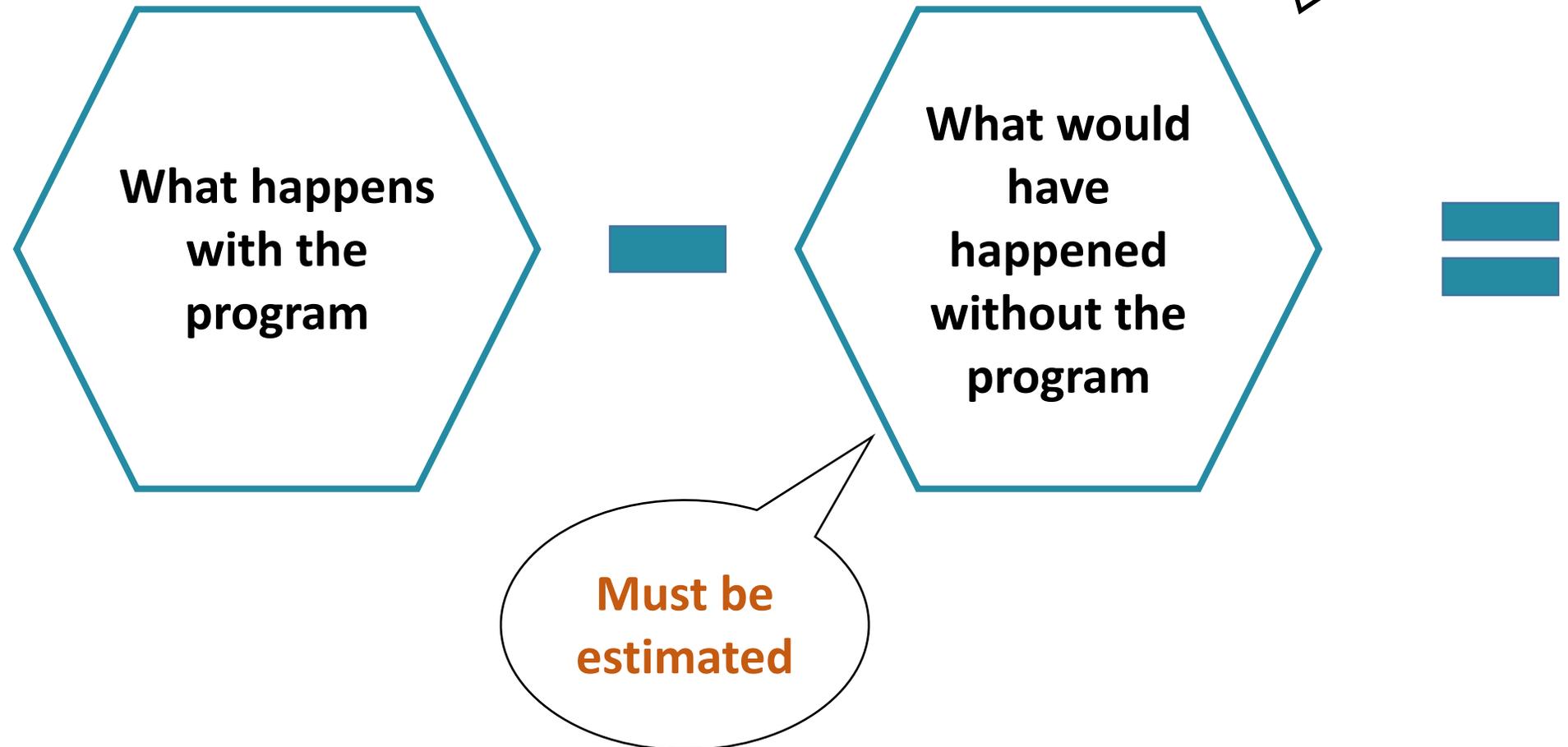
Impact Evaluation in Practice, second edition

How do we know that the program *caused* the observed effects?

What would have happened without the program?

The counterfactual

Program impact



I
M
P
A
C
T

**How can we estimate the
counterfactual?**

Some options to consider

- Before-after comparisons
- Comparing participants and non-participants
- Randomized experiment

After-school reading program

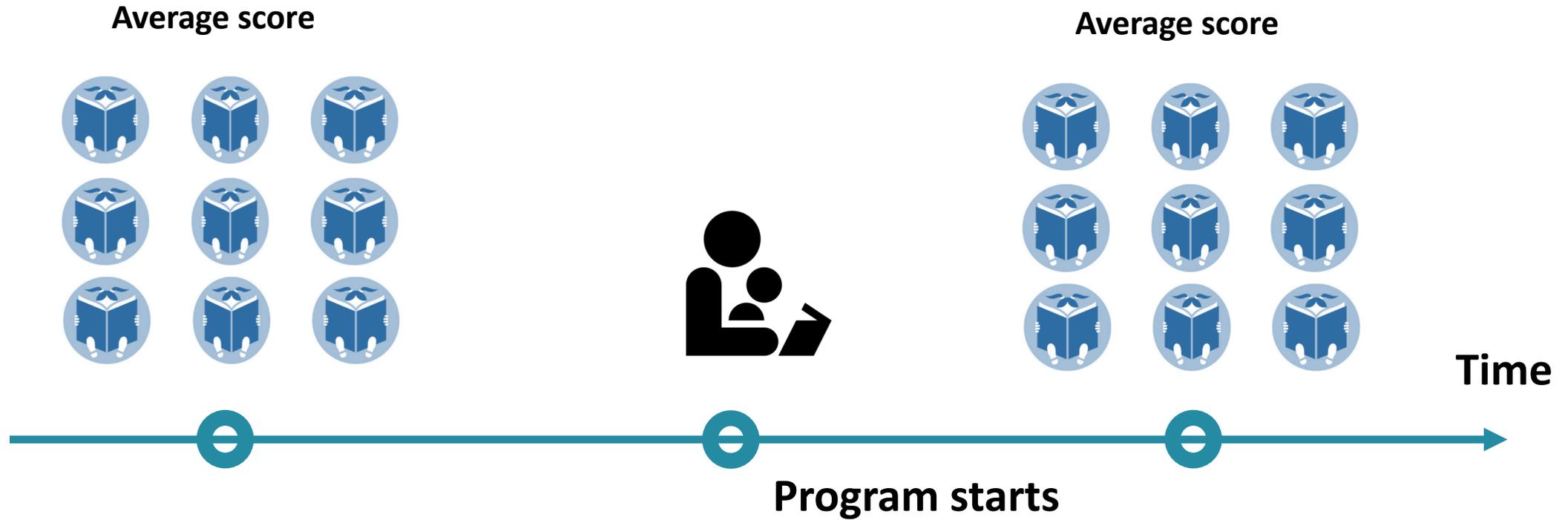




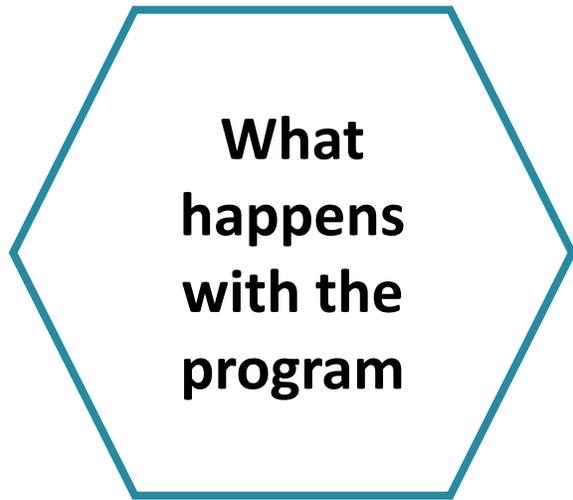
Time



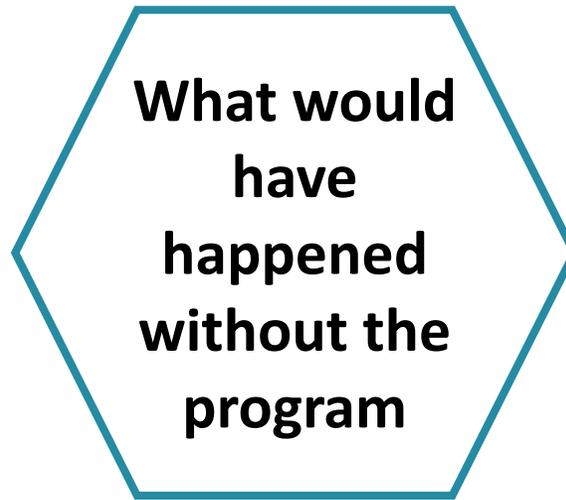
Program starts



Before-after comparisons



Average score after program



Average score before program



I
M
P
A
C
T



???

Before-after comparisons

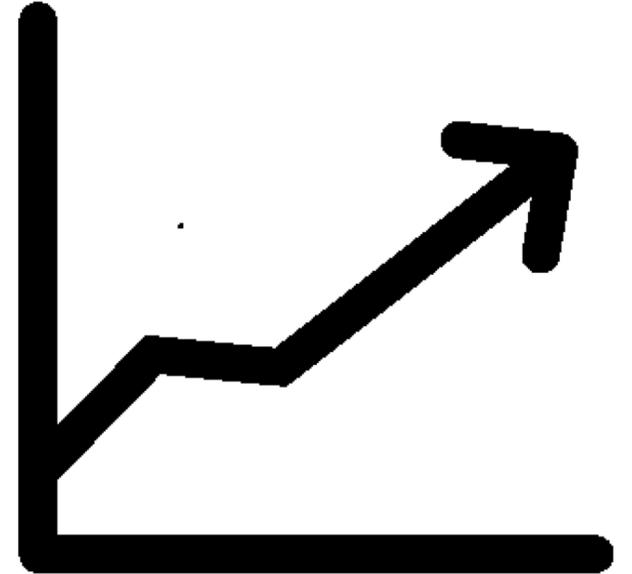
Many other things might have happened over time



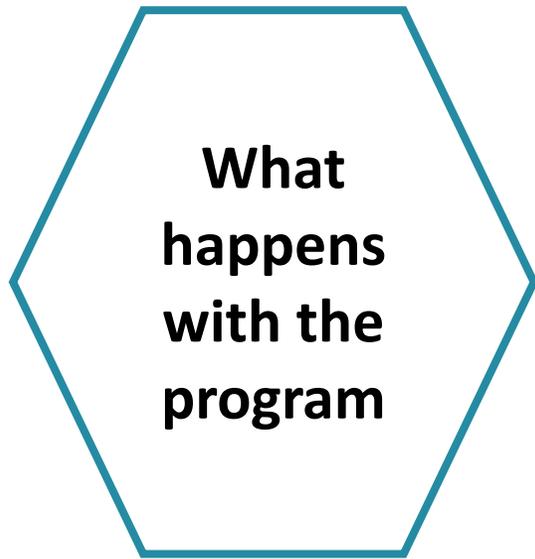
Learning during the school day



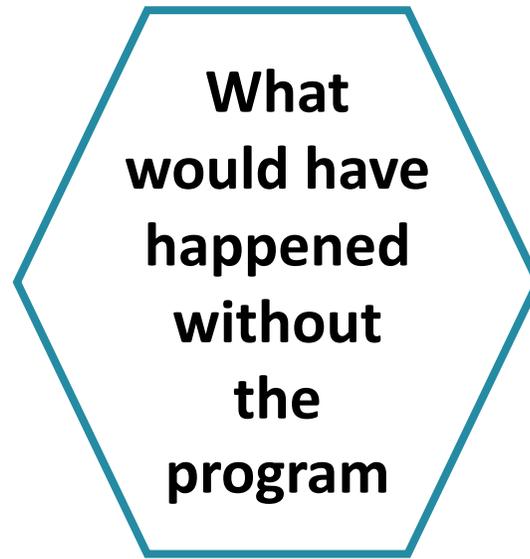
Other child-targeted activities



Income fluctuations



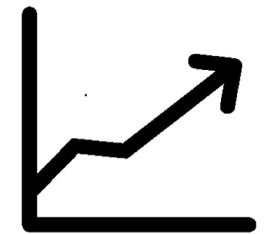
Average score after program



Average score before program



I
M
P
A
C
T



Before-after comparisons

Before-after comparisons give us a biased estimate of impact



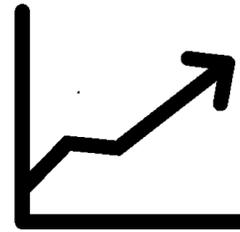
+



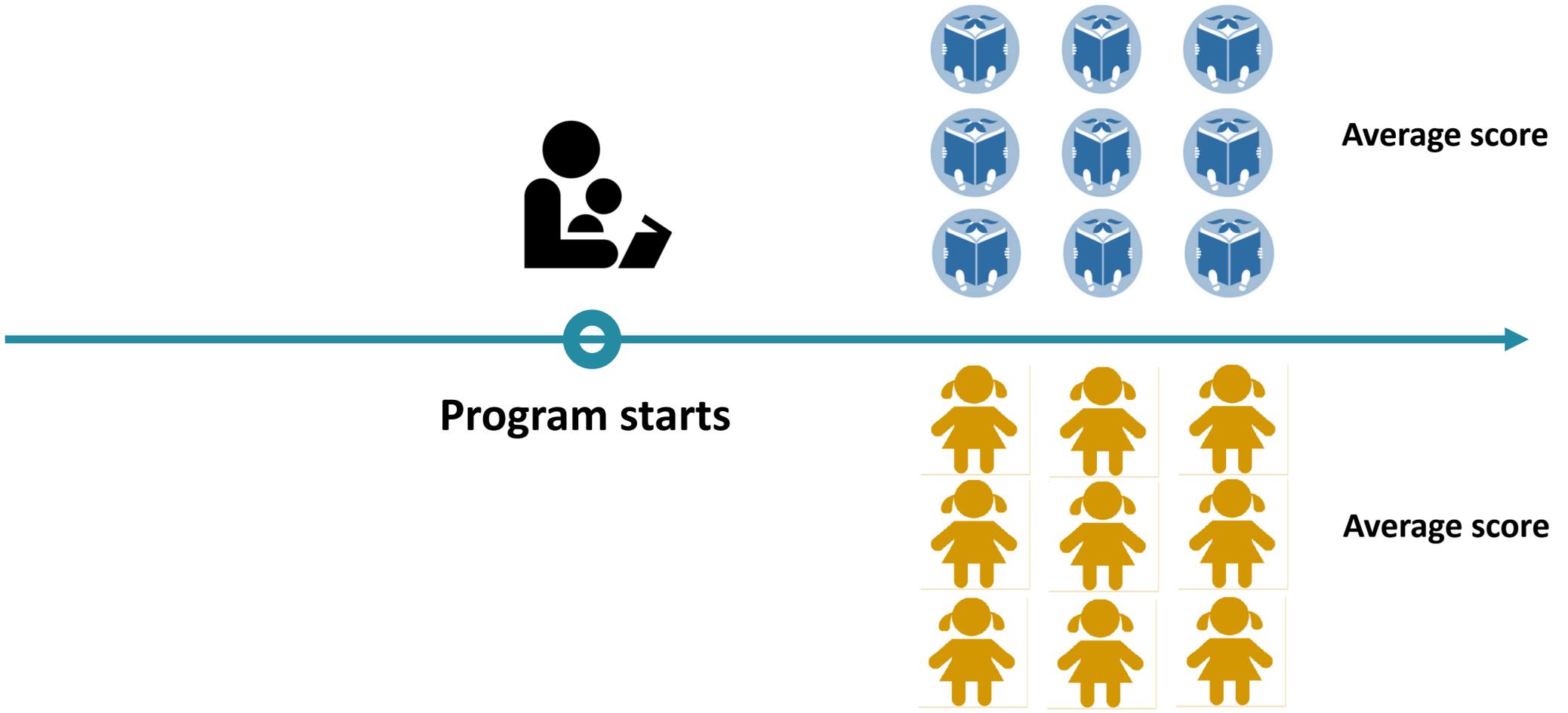
+



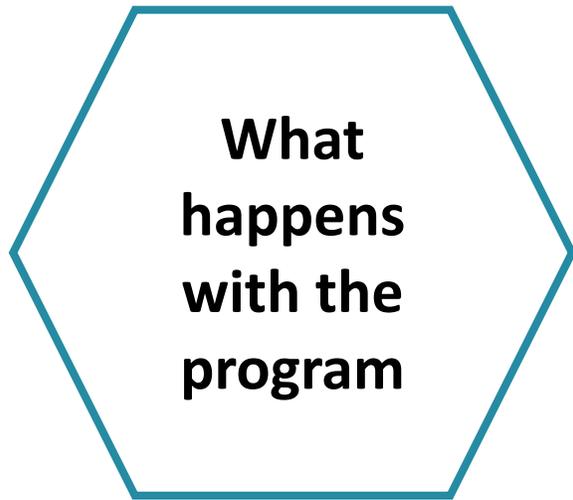
+



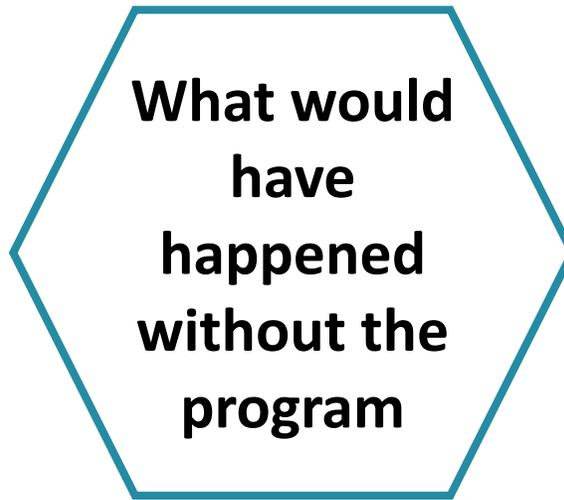
BIAS



Comparing participants and non-participants



Average score of participants



Average score of non-participants



I
M
P
A
C
T

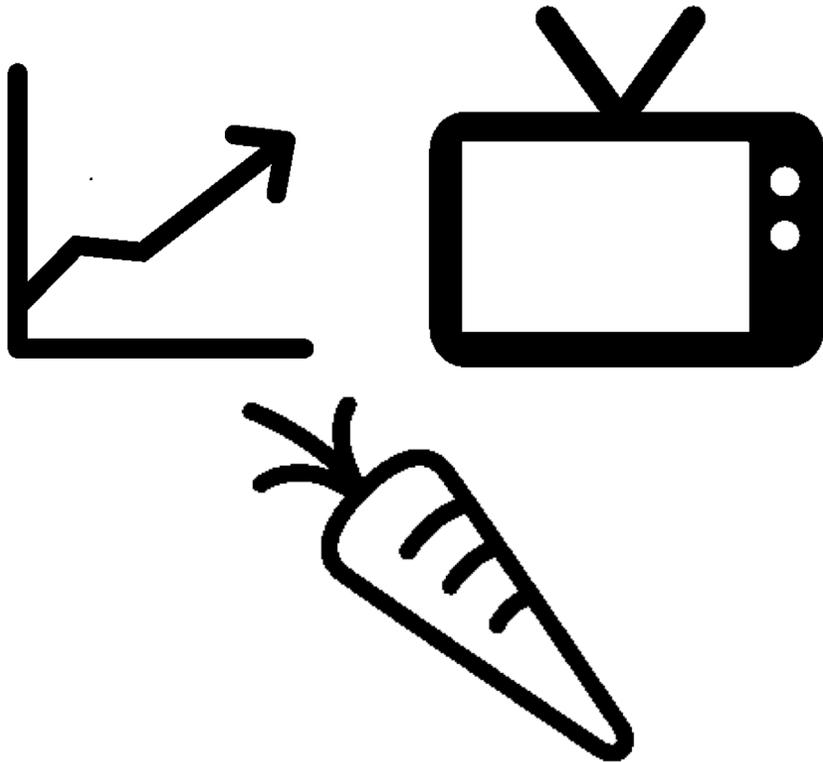


???

Comparing participants and non-participants

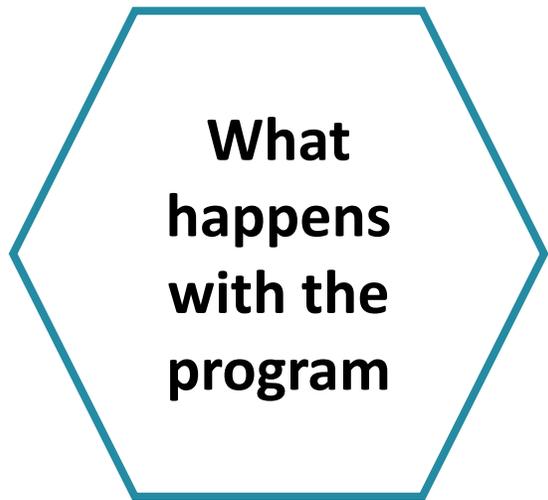
Participants and non-participants can differ along many dimensions

Observable

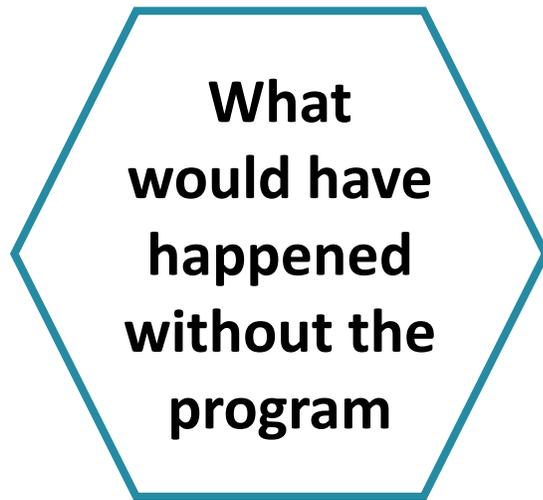


Unobservable





Average score of participants



Average score of non-participants

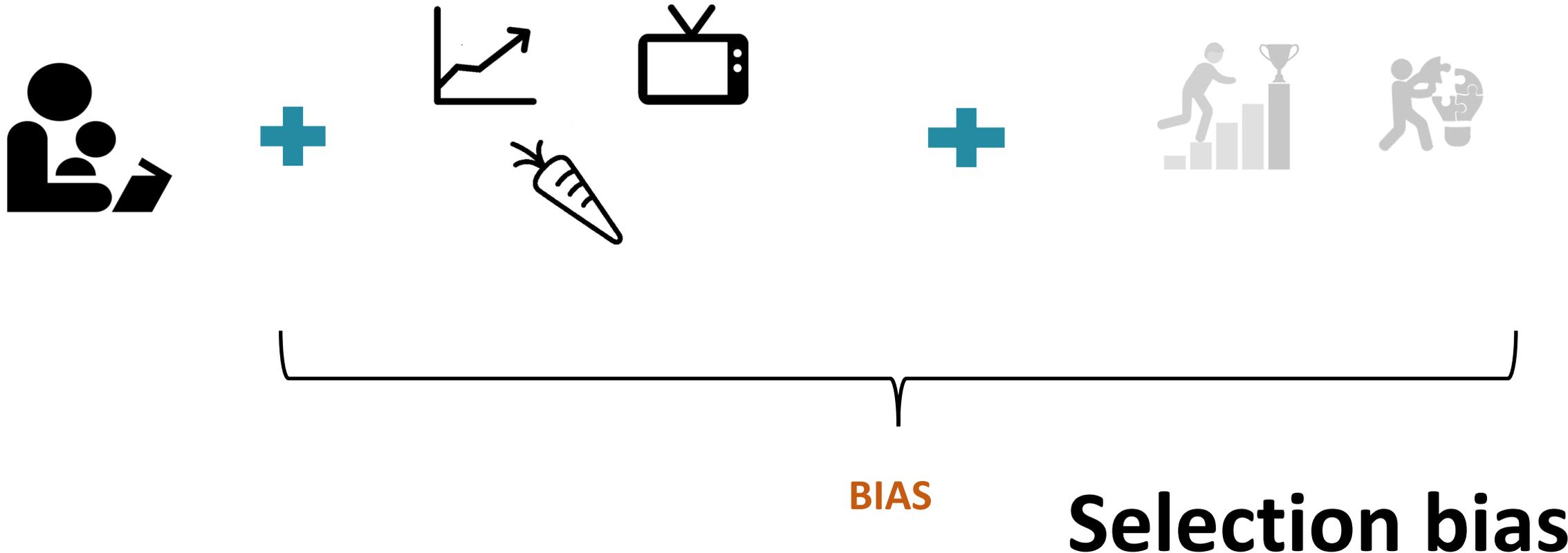


I
M
P
A
C
T

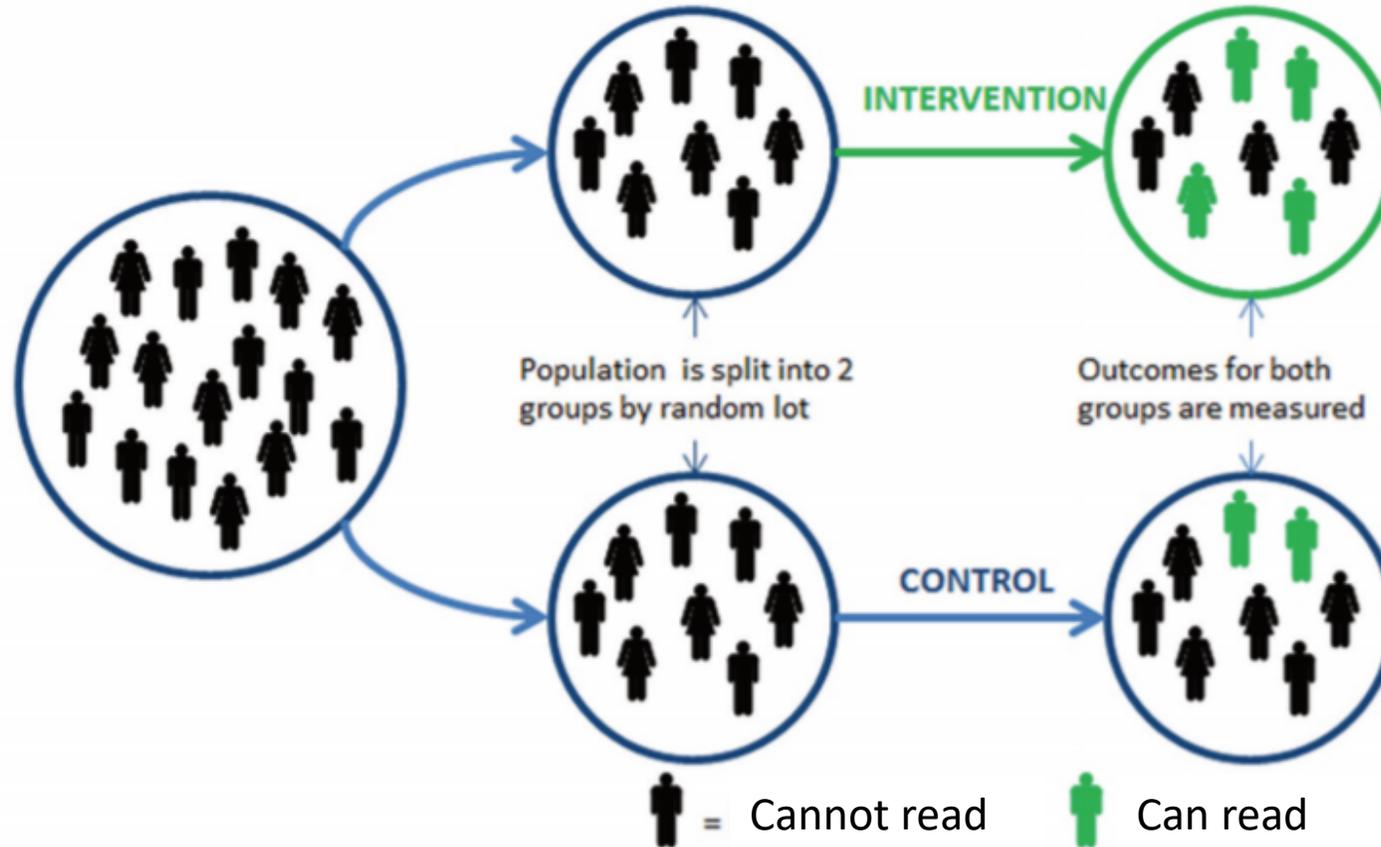


Comparing participants and non-participants

Comparisons of participants and non-participants give us a biased estimate of impact



Can experiments help?



There are many different types of children



There's even variation within type



For each child, let's flip a coin.

Heads → Treatment

Tails → Control

For each child, let's flip a coin.

**Every child has an equal chance
of being assigned to treatment**

Randomized assignment ensures balance

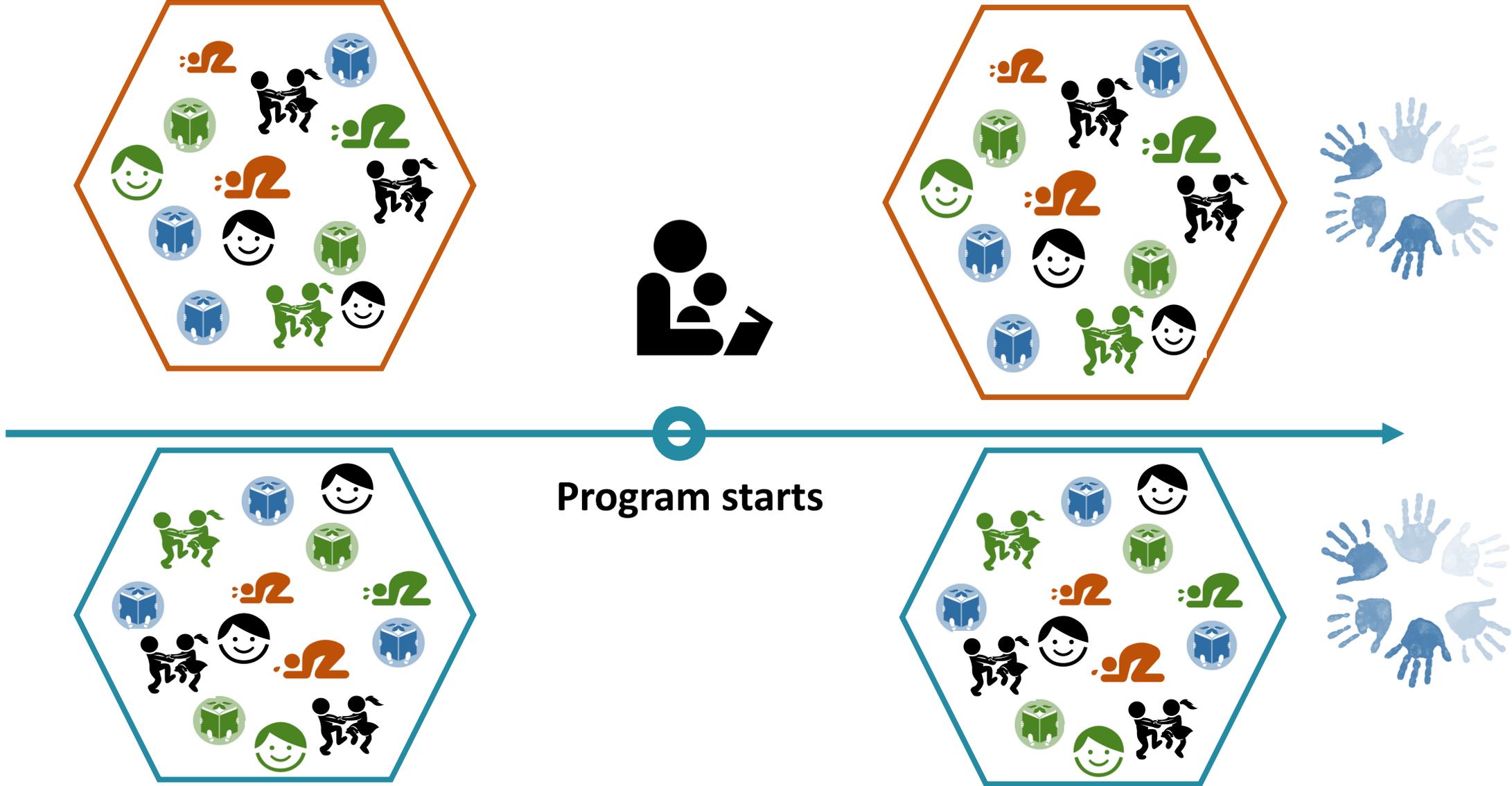


TREATMENT: Gets program

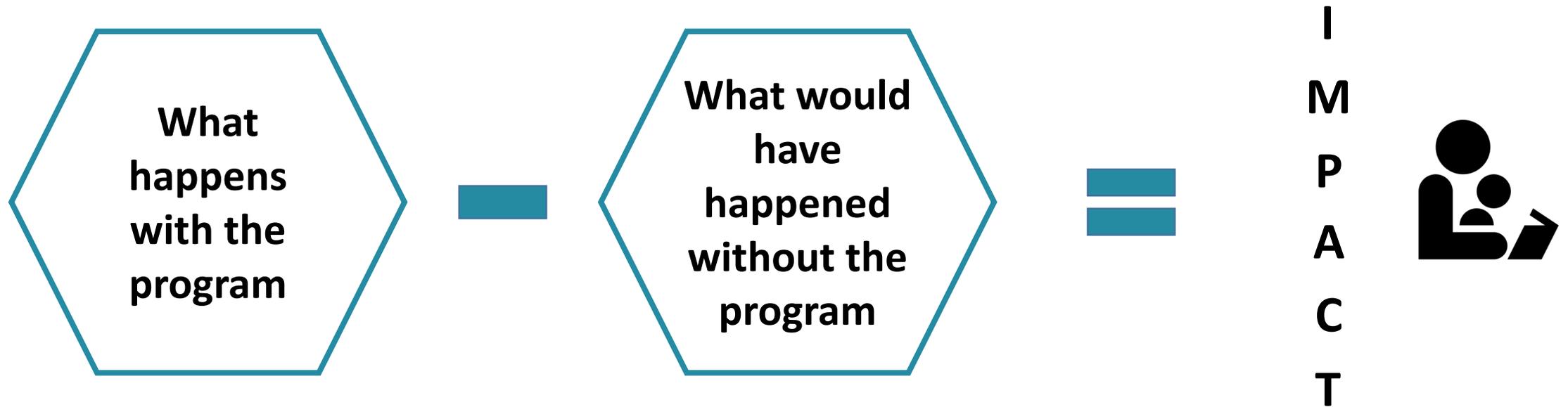


CONTROL: Does not get program

Time varying elements also balanced



Randomized assignment helps isolate impact

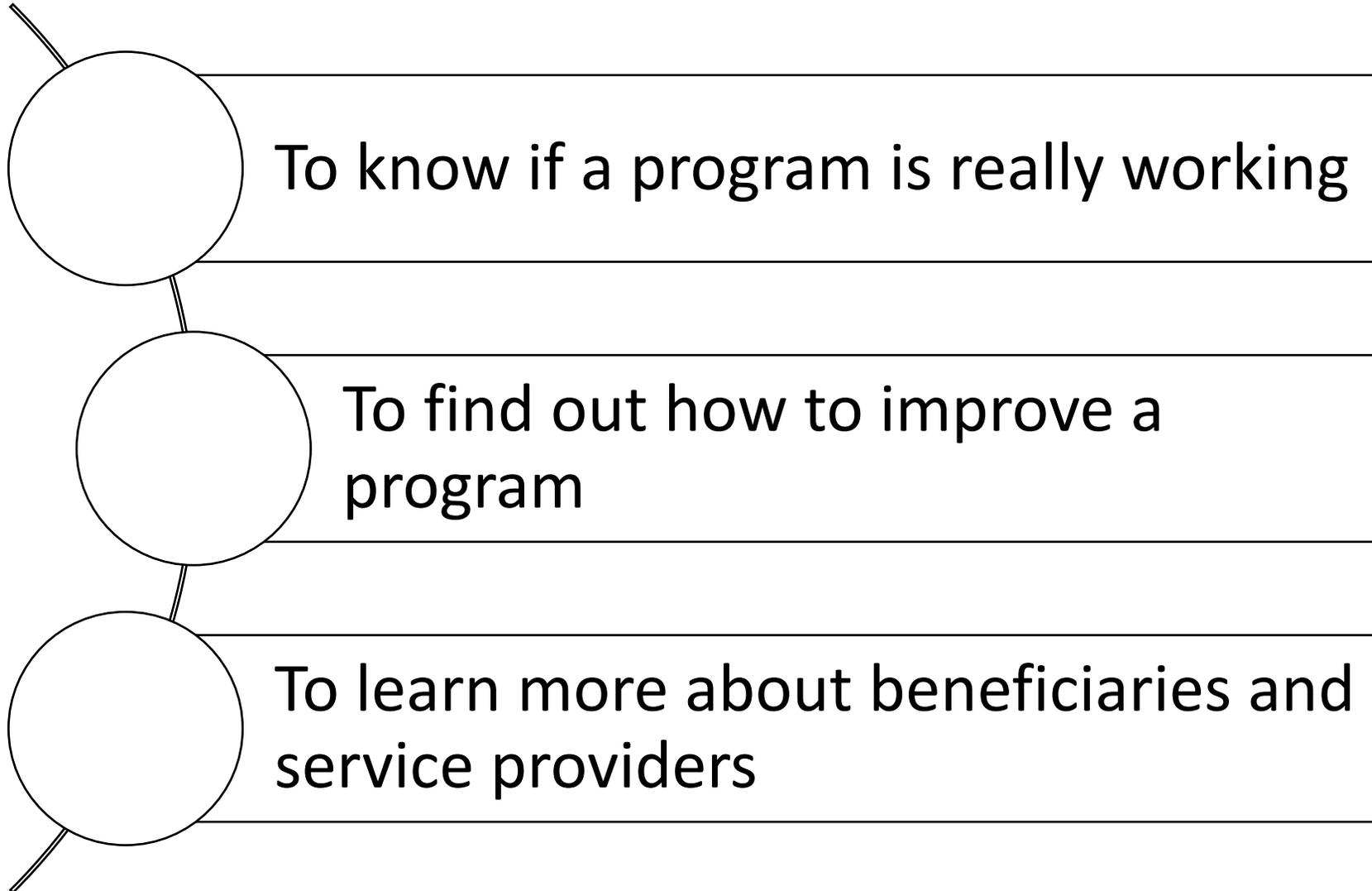


Average score of TREATMENT after program

Average score of CONTROL after program

Average treatment effect

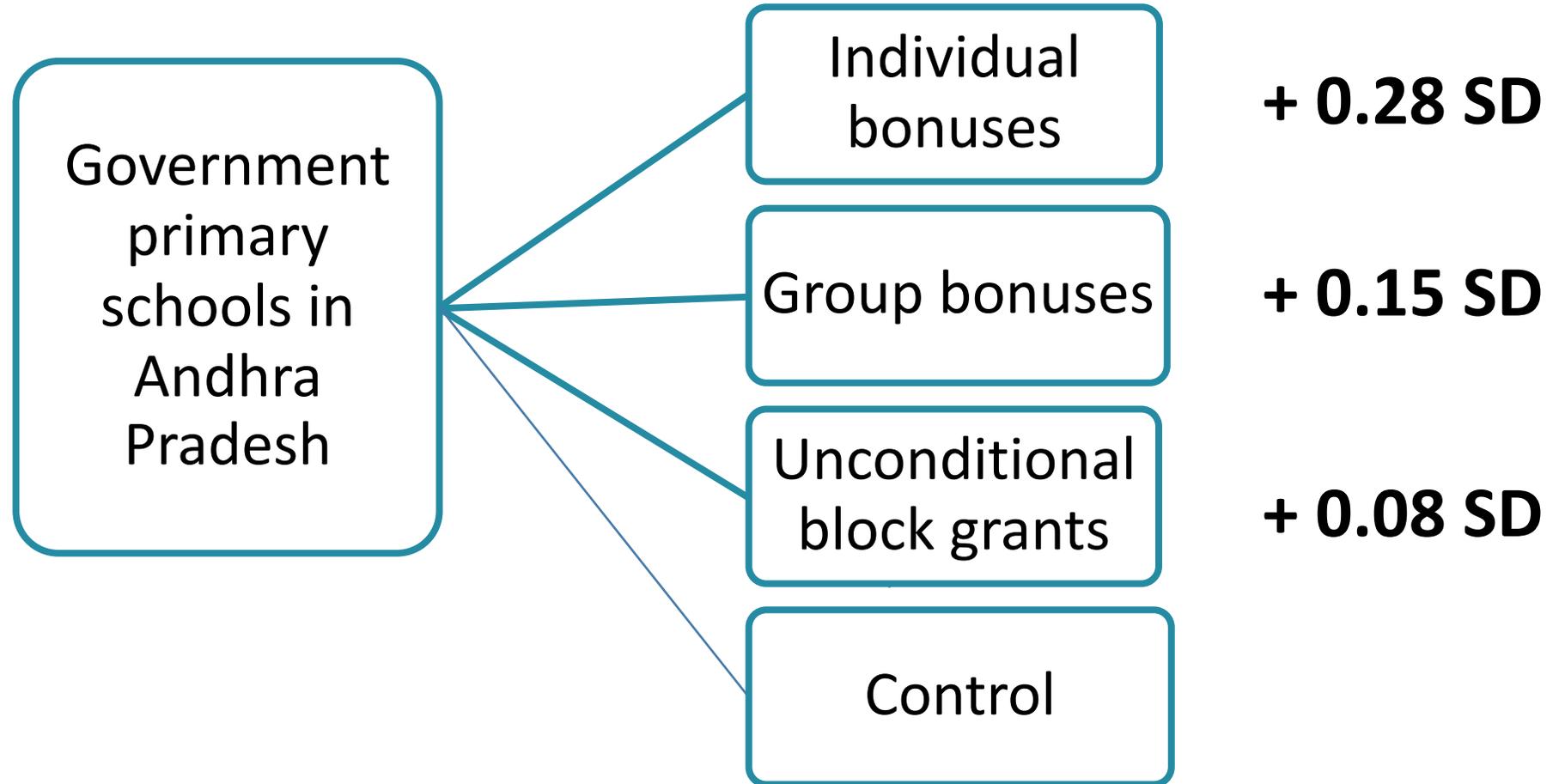
Why should we do experiments?



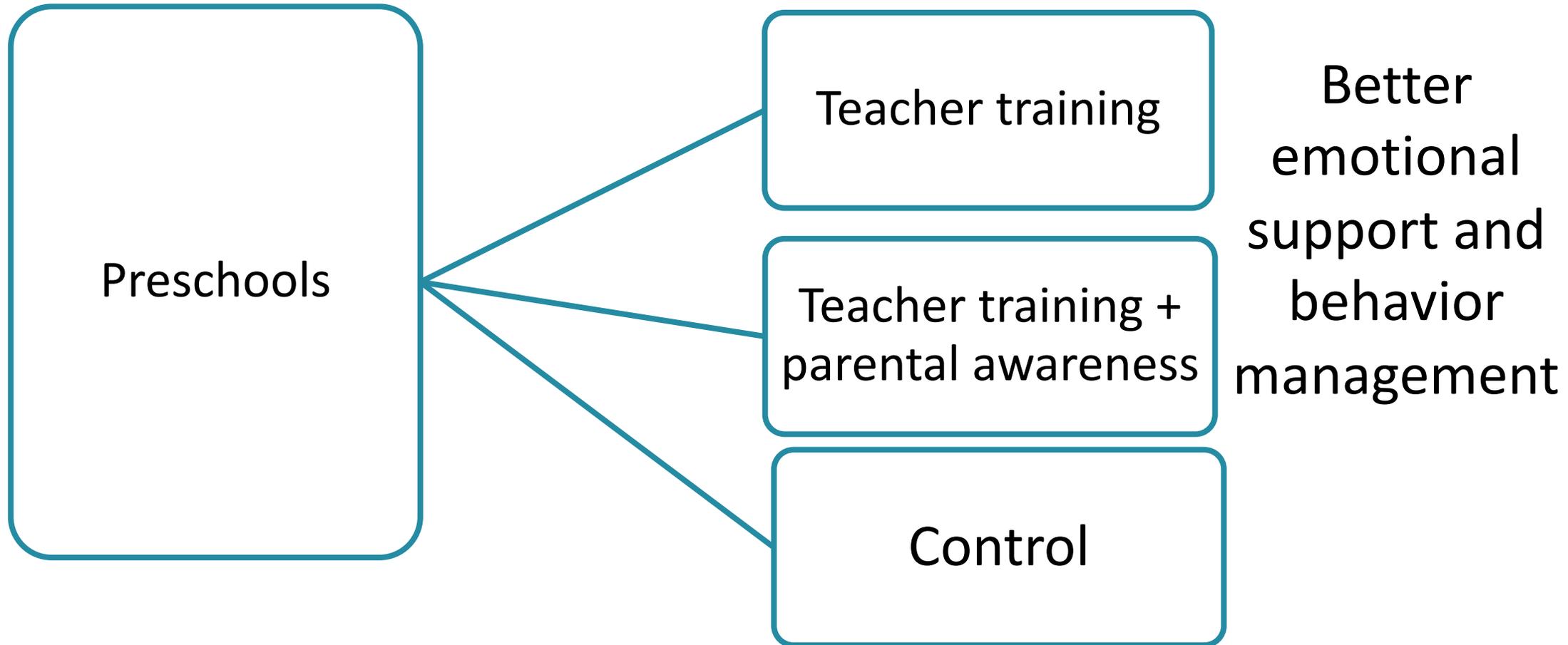


Testing whether programs work

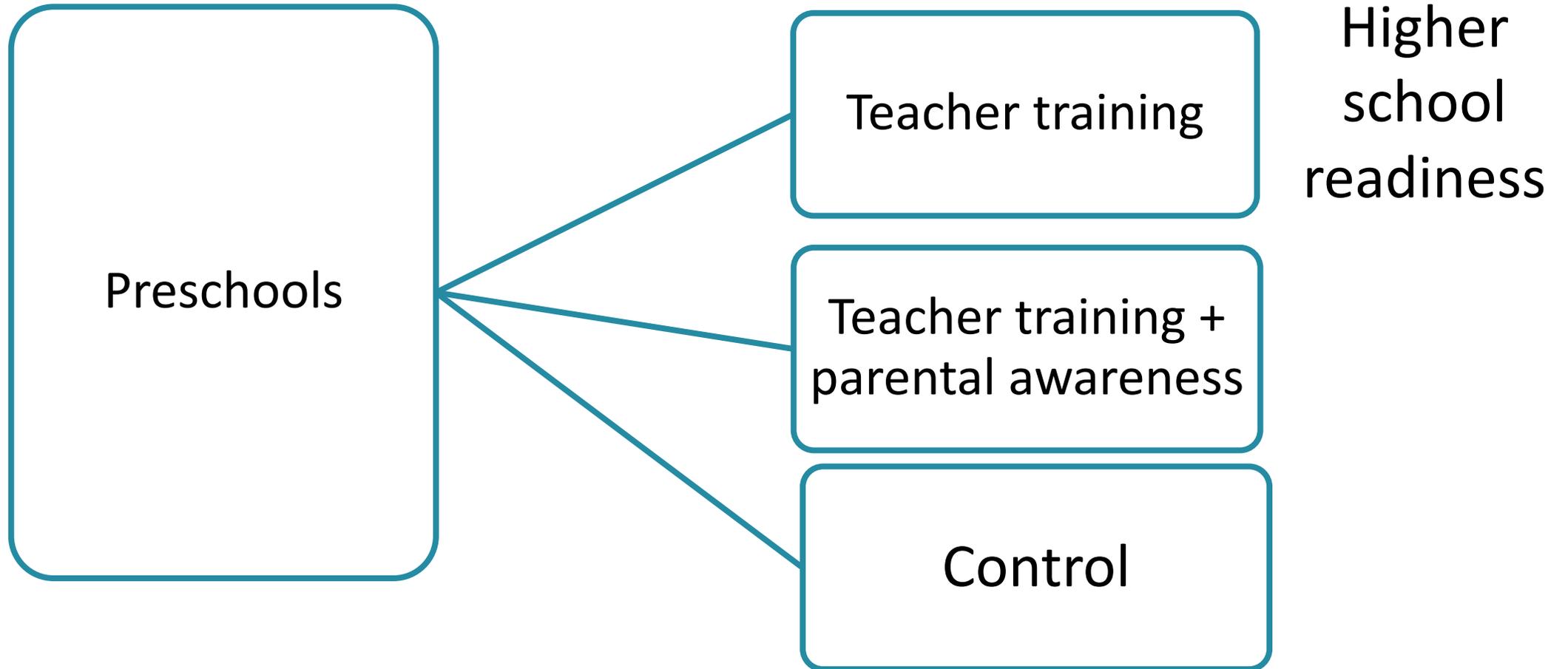
Teacher performance pay in India



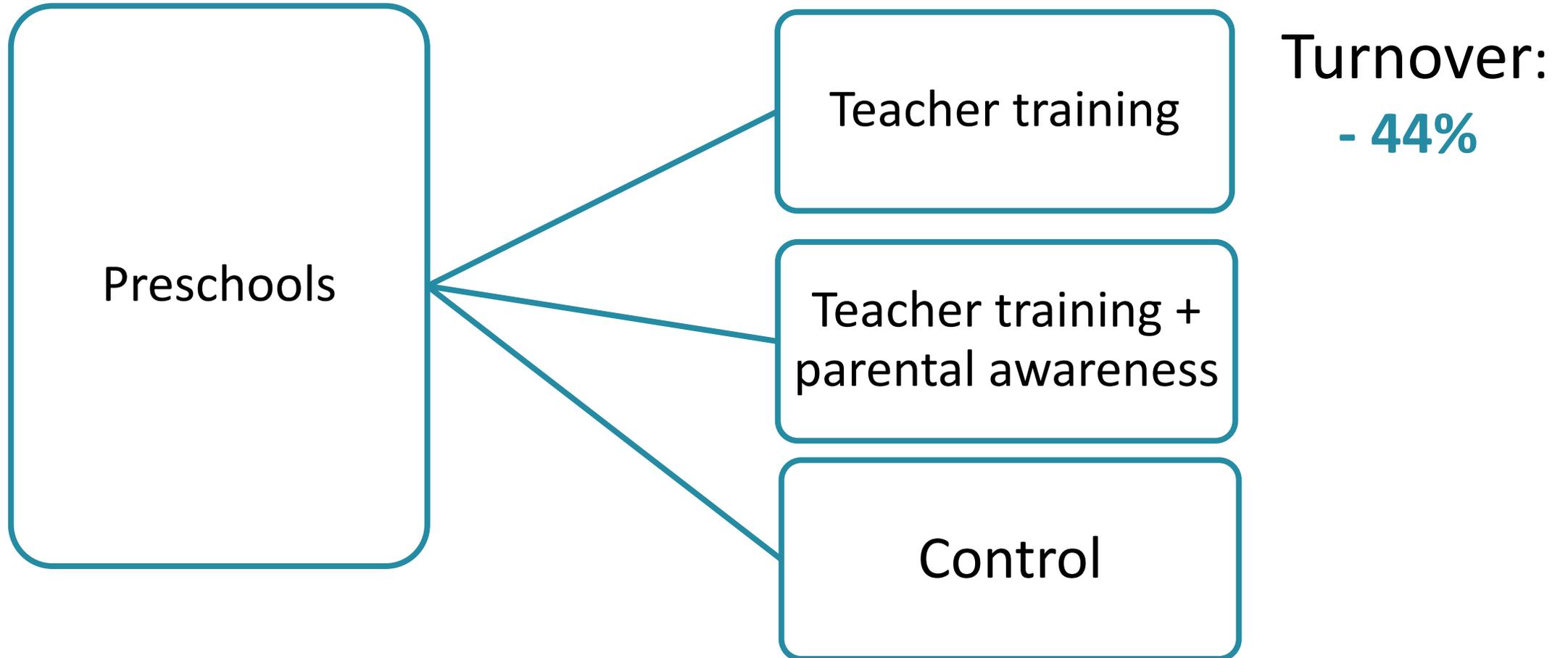
Improving pre-primary school quality in Ghana



Improving pre-primary school quality in Ghana



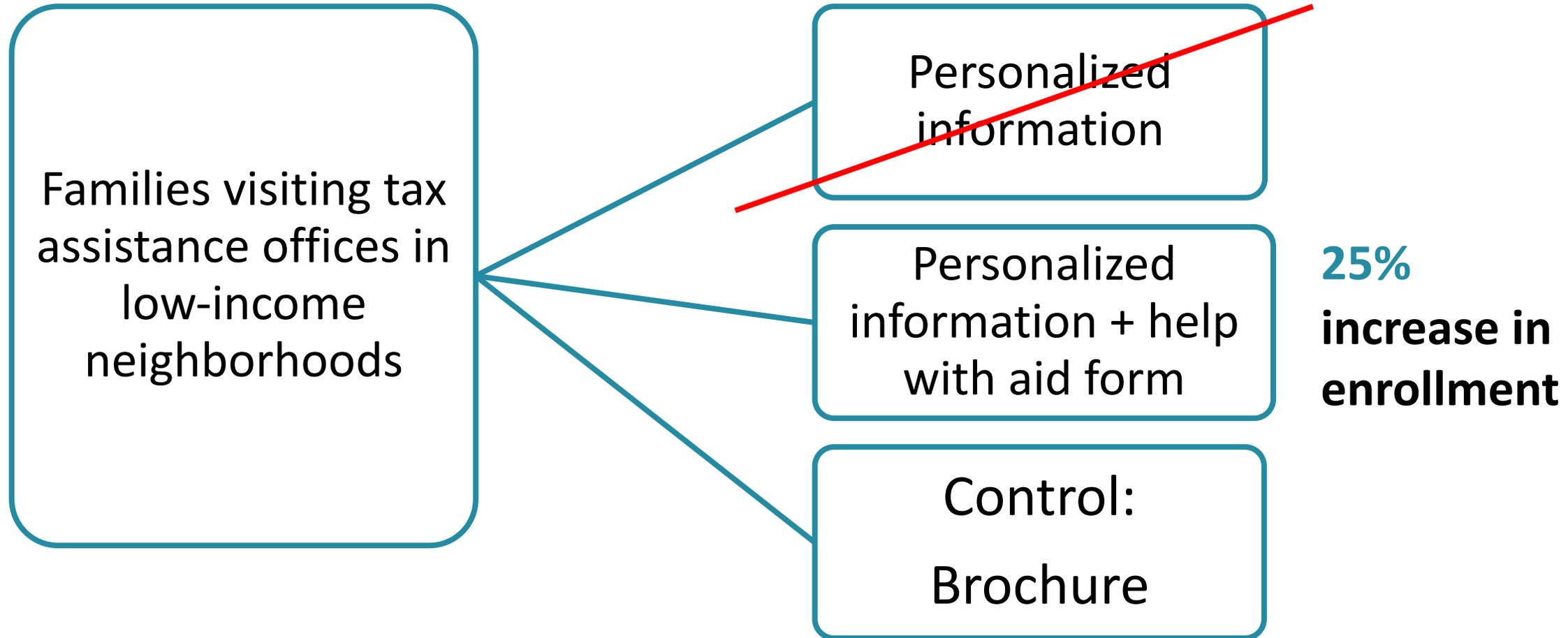
Improving pre-primary school quality in Ghana



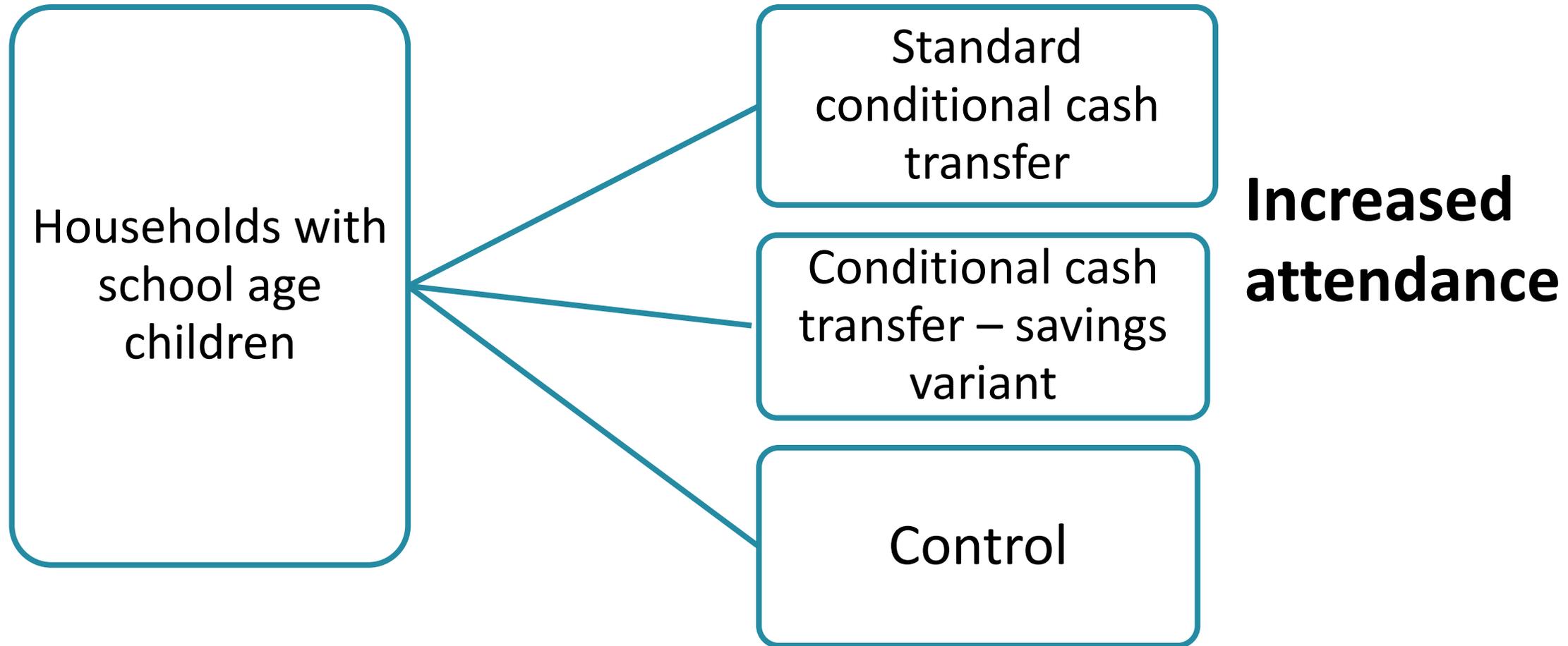


Improving program design

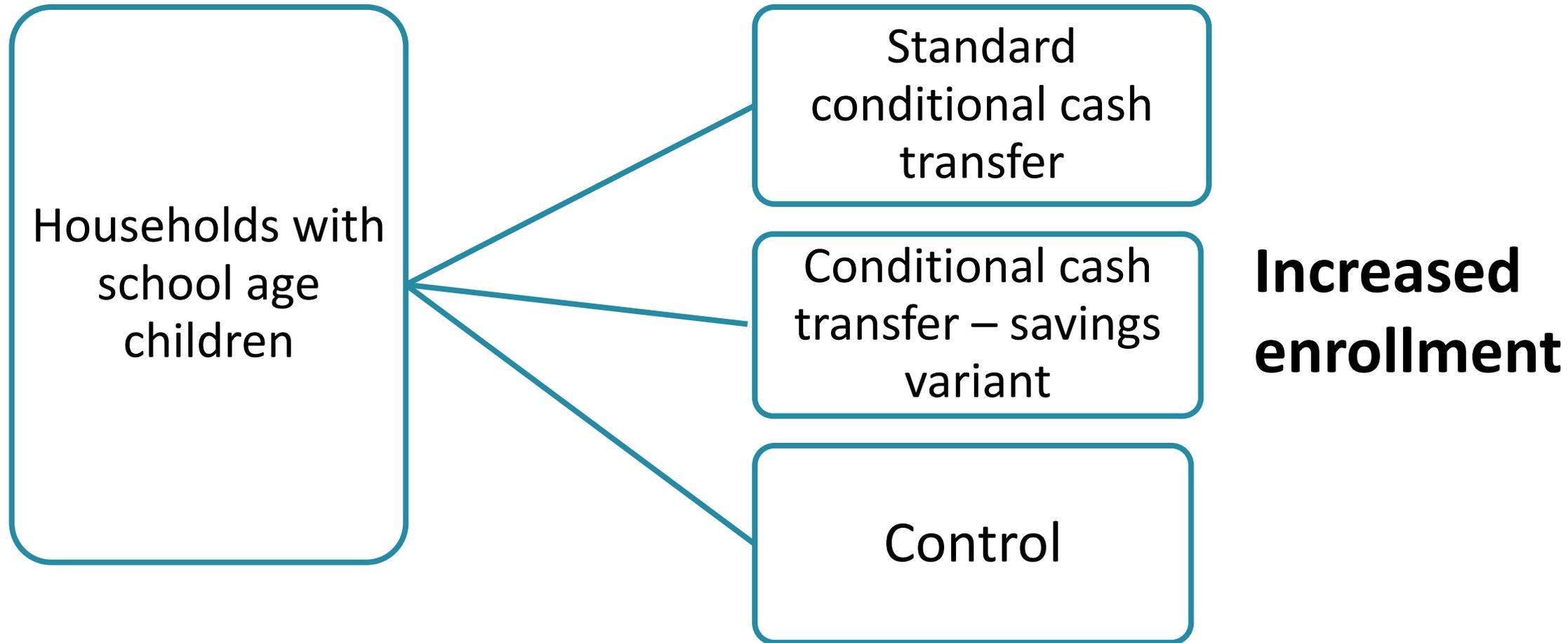
Increasing college attendance in the US



Increasing secondary school enrolment in Colombia



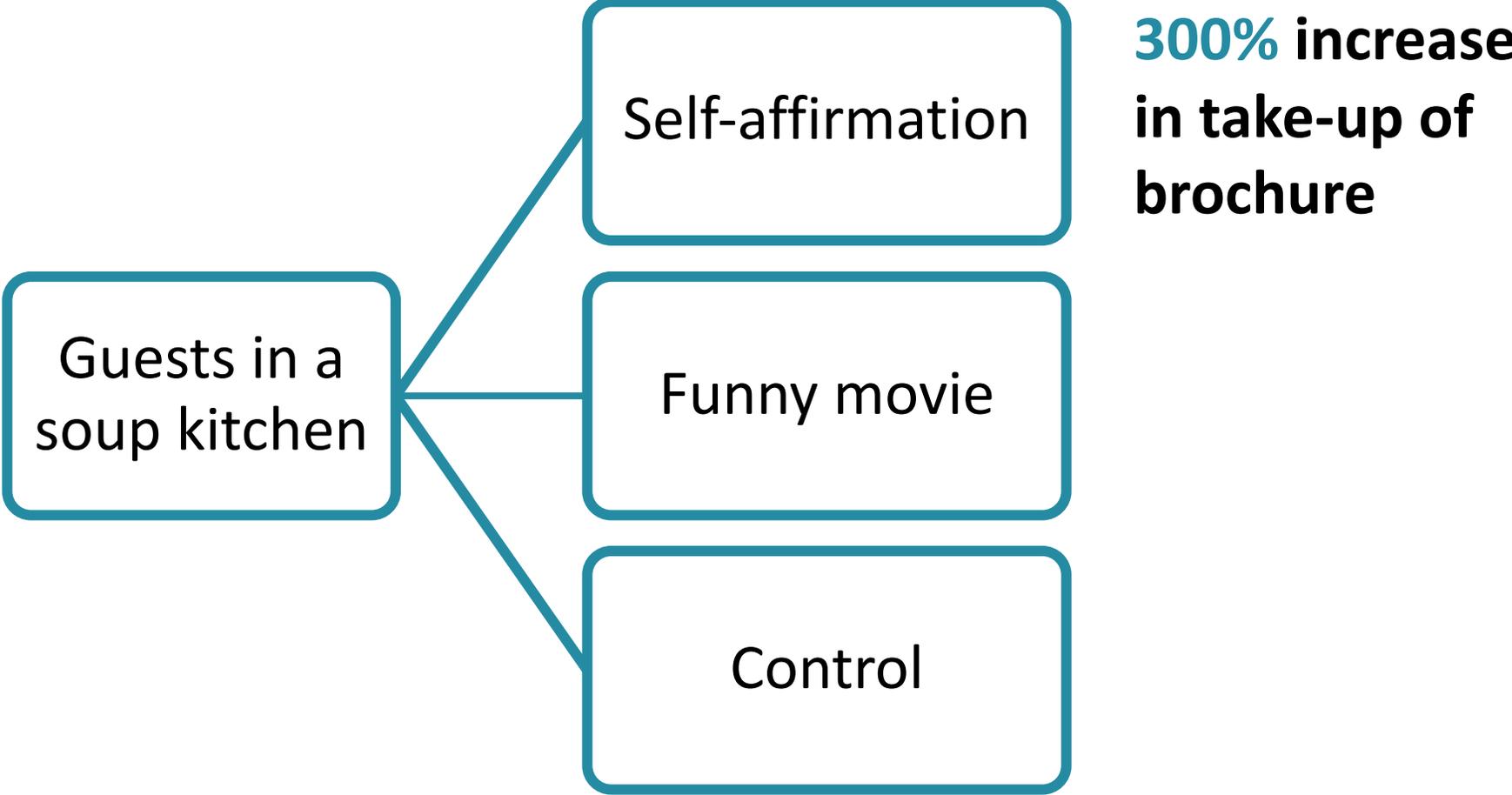
Increasing secondary school enrolment in Colombia





**Learning about beneficiaries or
service providers**

Improving interest in anti-poverty programs in the US





Strategic Impact
Evaluation Fund

<http://www.worldbank.org/sief>

THANK YOU