

ABDUL LATIF JAMEEL
Poverty Action Lab

TRANSLATING RESEARCH INTO ACTION

**Globally informed, locally
grounded policy advice**

Rachel Glennerster
J-PAL

Course Overview



1. What is Evaluation?
2. Outcomes, Impact, and Indicators
3. Why Randomize?
4. How to Randomize
5. Sampling and Sample Size
6. Threats and Analysis
- 7. Generalizability**
8. Project from Start to Finish

The challenge of using evidence

- Dramatic rise in the number of rigorous impact evaluations in developing countries in last 20 years
- Unlikely to be rigorous evaluation of precisely the program policy makers wants to introduce in exactly same location
- How should we respond?
 - Wait to act until there is more evidence?
 - Always do new rigorous eval before introducing in new context
 - Only use less rigorous local evidence?
 - Use results from study in another context?
 - Only use from other countries if at least X replications or if replicated in a similar enough context?

The challenge of using evidence II

- Rigorous impact evaluations are hard to do well and we underutilize their potential if we only learn about the precise program and context they evaluate
- But understanding local needs, and informal and formal institutions is critical to good policy
- We should do more replications of RCTs of similar programs in different contexts but there are limits
- Policy makers never have 100% certainty
 - Basu (2014) tomorrow is a new context
 - Is imperfect evidence **likely** to be worse than no global evidence?

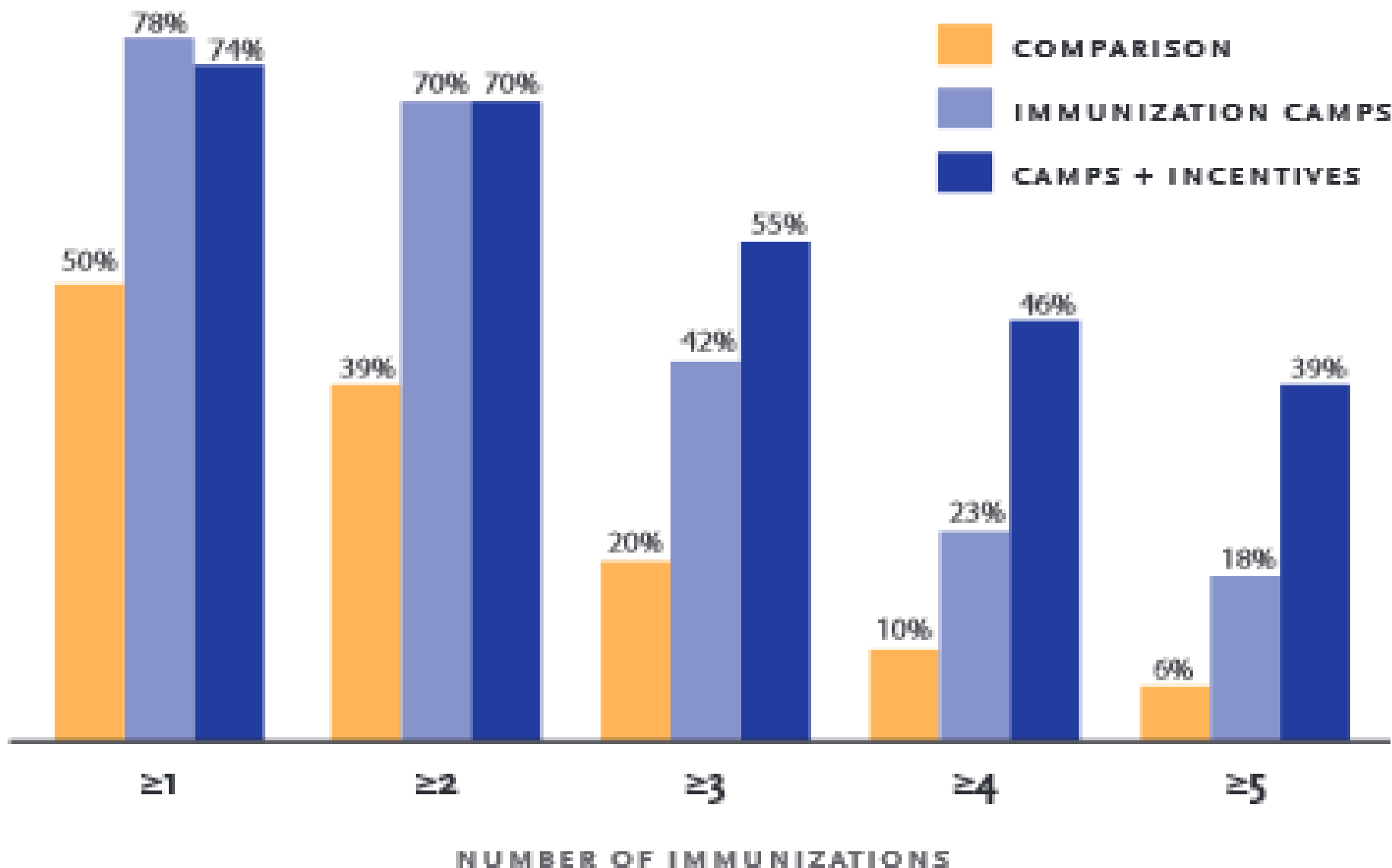
Overview of theory approach to evidence

- Evidence from a single RCT is only one part of the puzzle
- We use it to adjust our “priors” which are based on theory, descriptive work, other empirical evidence
- Putting evidence into a theoretical overview allows more efficient use of different forms of evidence than “black box”
 - allows us to be more precise about what a “similar context” is
- E.g. on improving immunization in a West African country
- Draw on a theory based overview of 70+ RCTs on health econ in dev countries (Kremer and Glennerster, 2012)

Non cash incentives for immunization in Rajasthan

- Seva Mandir program to increase immunization rates in rural Rajasthan, tested with RCT
 - Banerjee, Duflo, Glennerster, Kothari, 2010
- Fixing supply: regular monthly immunization camps with nurse present without fail
- Incentive: 1kg dahl for every vaccination, set of plates on completed immunization schedule

FIGURE 1: NUMBER OF IMMUNIZATIONS RECEIVED BY CHILDREN AGED 1-3 YEARS



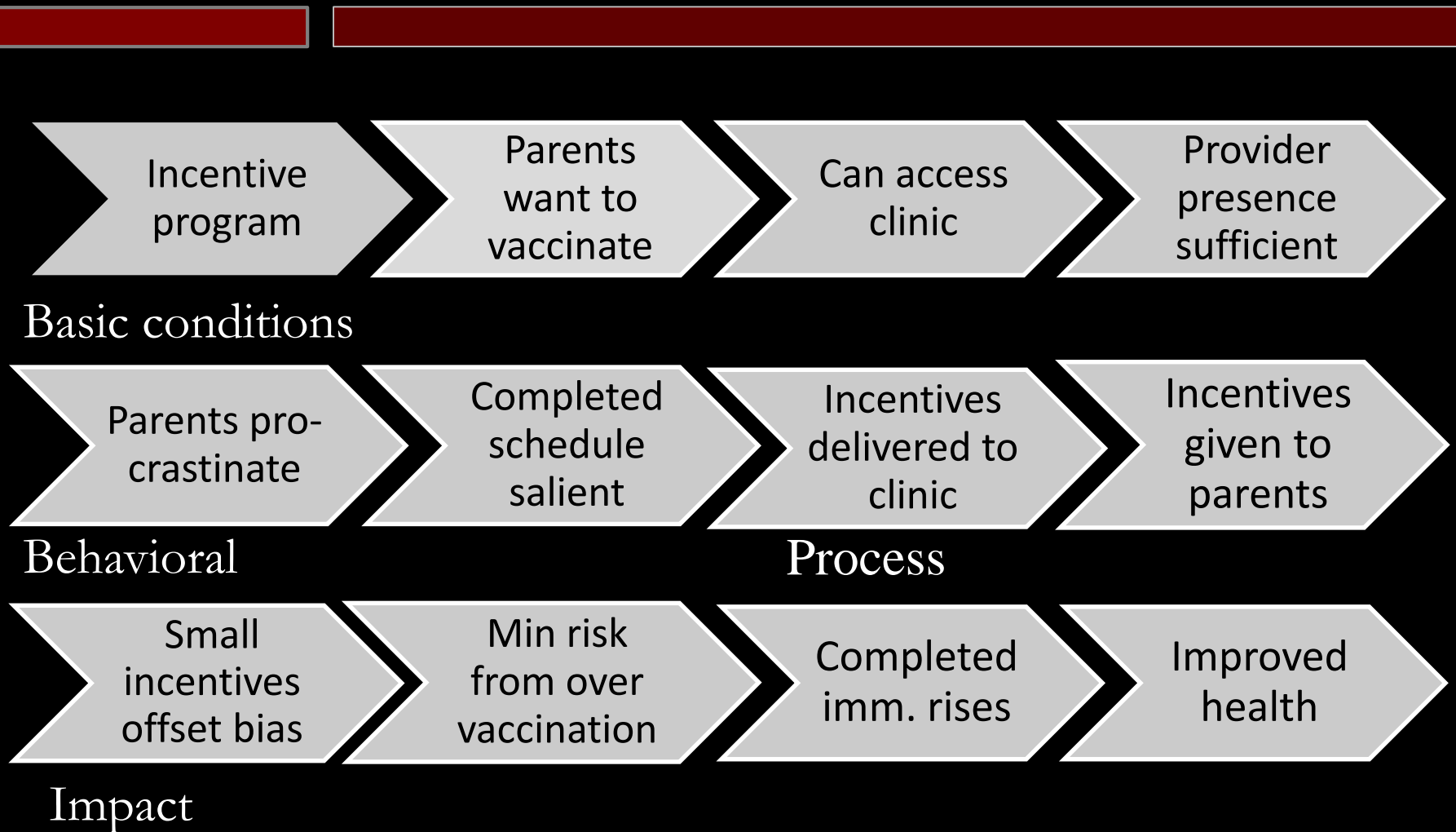
The “black box” approach to evidence

- If Govt in West Africa wanted to improved immunization rate, should they consider noncash incentives?
- What is our evidence of the following relationship?



- Only one RCT in South Asia not Africa
- Program conducted by NGO not government

Theory of change: incentives for immunization



Evidence on the basic conditions

- What evidence do we have on basic conditions?
 - Do parents want to immunize?
 - Is access to clinics adequate?
 - How big a barrier is health worker absenteeism?
- Descriptive evidence:
 - 54% of households within 1 hour walk of clinic
 - Health worker absenteeism 44%,
 - 84% of children receive DPT1
- Institutional knowledge:
 - unlike India, clinics often have multiple workers, only closed 12%.
Immunizations on specific days when absenteeism is lower

Take-up rates particularly informative

Immunization rates by antigen

	Country 1	Country 2
DPT1	84	47
DPT3	74	41
Measles	67	41
Fully immunized	49	38

Evidence on behavioral linkages in TOC

- People procrastinate and find it hard to stick with behavior they believe is good for them and their children
 - Good theoretical work showing how this small changes to a standard discounting model produces series of testable conclusions and can explain many stylized facts (e.g. Laibson, 1997)
 - Small changes in price of preventative products sharply reduces take up (9+ RCTs)
 - People are willing to pay to tie their own hands with commitment savings products: difficult to explain unless people know they are present biased (e.g. Gine et al. 2010)

Evidence on behavioral linkages in TOC II

- Small incentives can have big impacts on behavior
 - 30+ RCTs of CCTs but usually much bigger incentives (Fiszbein and Schady, 2009)
 - Malawi: smaller CCT same impact as bigger CCT (Baird et al 2010)
 - Small incentives for HIV testing (Thornton 2008 Malawi), age of marriage (Field et al, 2014 Bangladesh)
- Knowledge, or salience, of how many vaccinations are needed for completed immunization
 - Weaker evidence on the importance of salience
 - India study had different incentive at final vaccination: how important?

Evidence on process links in the ToC

- Process questions include:
 - Will the incentives be delivered regularly to the clinic?
 - Will the incentives be given to parents appropriately?
- Harder to generalize on these process questions from one program and country to another
- Not just a challenge from learning from RCTs, good implementation is a constant struggle in development
- This is why we need monitoring for every program:
 - We may be confident a program will work if it is delivered, but we need to make sure it is delivered appropriately

Can RCTs tell us about details of delivery?

- Yes, but harder than in other areas.
 - Very similar results across contexts for consumer behavior (68 RCTs)
 - More varied results on provider behavior (6 RCTs)
- Providers are humans too, why harder to predict?
 - Work in bureaucratic settings with complex incentives
 - Theory of change longer, with many more steps
 - Increasingly RCTs are breaking down these steps and testing them
 - Many fewer RCTs on provider behavior than consumer behavior
- Some delivery harder than others
 - Incentives through MPESA, or cell phone vouchers
 - Less concern about bureaucratic incentives

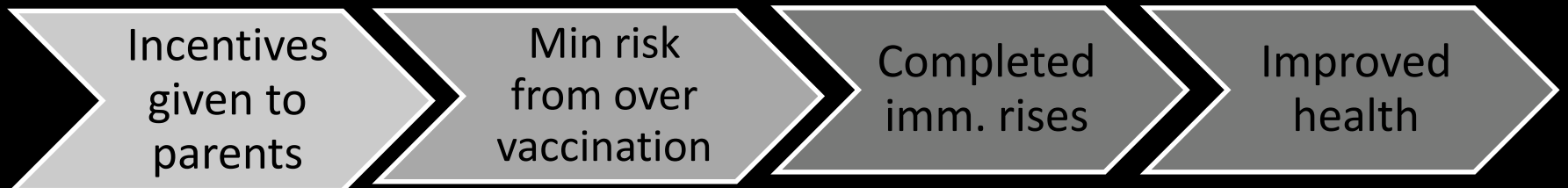
ToC incentives for immunization: Country 1



Local descriptive evidence suggest basic conditions hold



Strong RCT evidence of behavioral conditions Need to monitor process



If all else holds, expect to achieve

Defining “similar context”: sugar daddies

- Dupas (2011) info on relative HIV infection rates of young vs older men reduced teenage pregnancy with older men
 - girls trade off costs and ben of sex, older men give more gifts + protect if get pregnant. Don't know older men higher HIV risky
- Theory: how people respond to information depends on difference between new info and old info
- Rwandan government wants to scale the program. How do we measure if this is the right context?
 - Do older men have higher HIV rate than younger men?
 - Do girls know about the difference in relative risk?
 - What do they know about absolute risk?

Don't take discretion out of literature reviews

- Moving towards a medical model, increasing use in economics of synthetic reviews and meta-analyses
 - Good to avoid selective inclusion of studies, though need tight quality criteria for inclusion, more weight to better studies
- By averaging results across studies, meta-analyses often fail to pick up underlying broader lessons
 - In attempting to eliminate bias from judgment, we lose role theory
- Effective ed programs focus learning at level of the child
 - Meta analyses classified these studies under technology, class size, materials, teaching resources
 - Meta-analyses don't take into account local descriptive data

When do we stop evaluating?



- If we have enough evidence to act, do we have enough evidence to stop evaluating impact? (always monitor)
 - No: we often need to act even when evidence is thin
- Often a big overlap between when have enough evidence to launch big new initiative and when still worth evaluating
- Questions may remain about best way to implement
 - Billions spent on CCTs. Very strong evidence they work, but important evaluations on how to make more effective
- Trade-off of evidence in new areas, vs more on existing

Concluding thoughts

- Design research for generalizability
 - Theory based RCTs can be very useful for policy because ask particularly generalizable questions
- Policy making requires drawing on different kinds of evidence, but that does not mean all evidence is equal
- Implementation is hard: knowing a program will have impact if implemented is a good place to start
 - Logistics pilots important part of scaling up, even if no new RCT