



EVALUATIONS WITH IMPACT: DECISION-FOCUSED IMPACT EVALUATION AS A PRACTICAL POLICYMAKING TOOL

Dr. Neil Buddy Shah, CEO, IDinsight | 3ie Evidence Week | 04/16/2015

Agenda

- About IDinsight
- Main Argument
- Status Quo: Knowledge-Focused Evaluations
- New Paradigm: Decision-Focused Evaluations
- Appropriate Roles
- Realizing the New Paradigm

About IDinsight

- Non-profit organization that helps development managers generate and use rigorous evidence to improve their programs.
- Design and conduct rigorous impact evaluations to serve as decision-making tools for implementers
 - demand-driven and attentive to the priorities and constraints of implementers
 - only do IE if it will guide a specific decision
- Since 2012, have completed experimental evaluations across Africa and Asia in health, governance, education, sanitation, energy and finance

Agenda

- About IDinsight
- Main Argument
- Status Quo: Knowledge-Focused Evaluations
- New Paradigm: Decision-Focused Evaluations
- Appropriate Roles
- Realizing the New Paradigm

Main argument

- Cost-effectiveness lens
 - We should analyze impact evaluation as any other development intervention – how can we make it more cost-effective at improving social outcomes?
- Context-specific lens
 - Effectiveness of a given development intervention is highly context-dependent

Main argument

- Status quo:
 - Global learning agenda → Focus on building bodies of evidence and contributing to development theory
 - Led to emphasis on “**knowledge-focused evaluations**”
- Future of impact evaluation:
 - Recalibration between **global learning agenda** and **local solution finding systems**
 - Greater use of “**decision-focused evaluations**”
 - **Appropriate use of “knowledge-focused evaluations”** when objective is to inform development theory

Agenda

- About IDinsight
- Main Argument
- Status Quo: Knowledge-Focused Evaluations
- New Paradigm: Decision-Focused Evaluations
- Appropriate Roles
- Realizing the New Paradigm

“Knowledge-focused evaluations”

- Definition:
 - Evaluations that are designed with a primary goal of contributing to a “global body of evidence” or “development theory”, usually through publication in an academic journal.
- Characteristics:
 - Researcher-driven
 - Prioritize internal validity in evaluation design
 - Geared towards publication in an academic, peer-reviewed journal
 - Relatively long timeframe
 - Larger budgets

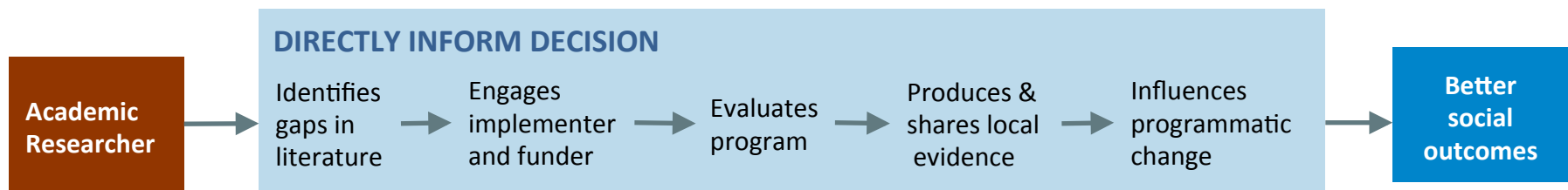
Successes of knowledge-focused evaluations

- Multi-country scale-up of cost-effective interventions
 - deworming
 - chlorine dispensers
- Deflating over-hyped interventions
 - microlending
 - one laptop per child
- More sophisticated discourse on impact and evidence
- Development theory advancements

Knowledge-focused evaluation theory of change

Channels

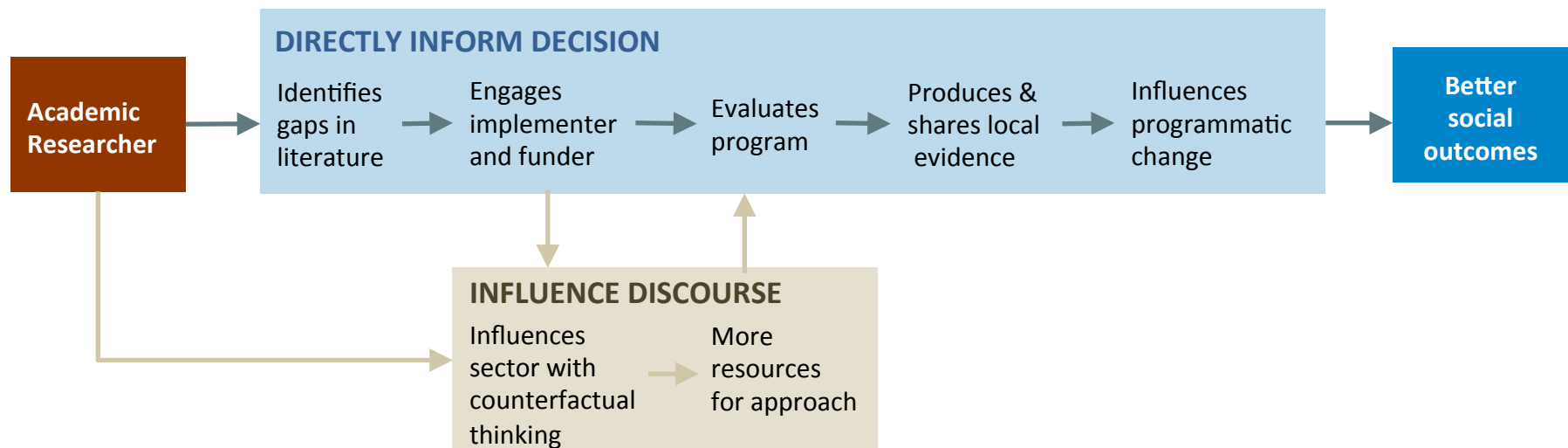
1. Directly inform implementing partner's decision



Knowledge-focused evaluation theory of change

Channels

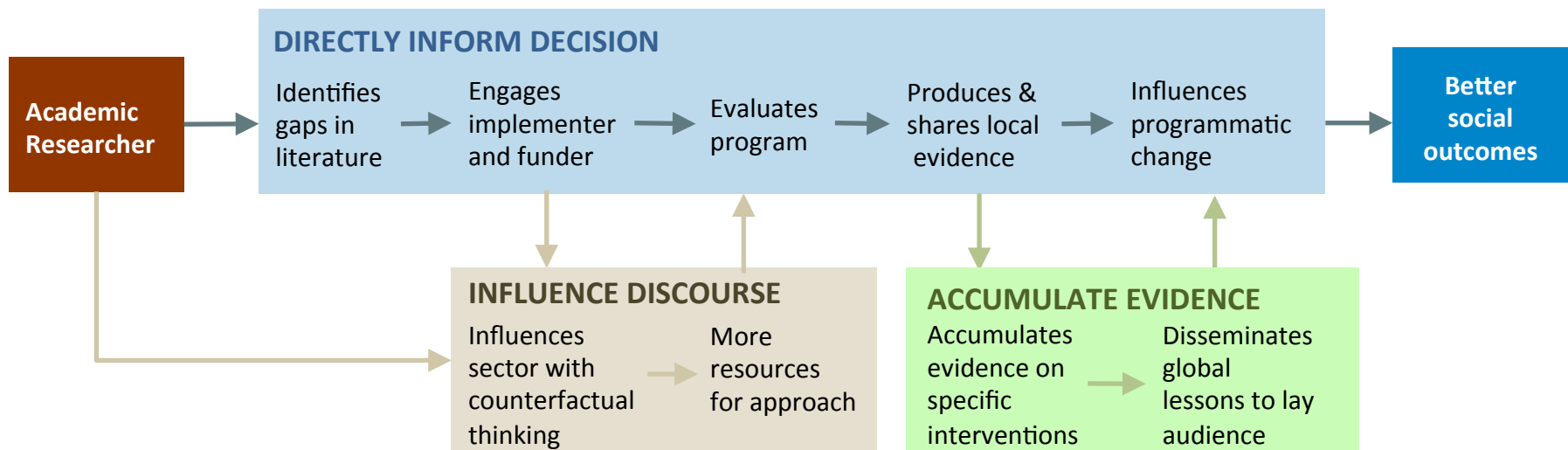
1. Directly inform implementing partner's decision
2. Influence development discourse



Knowledge-focused evaluation theory of change

Channels

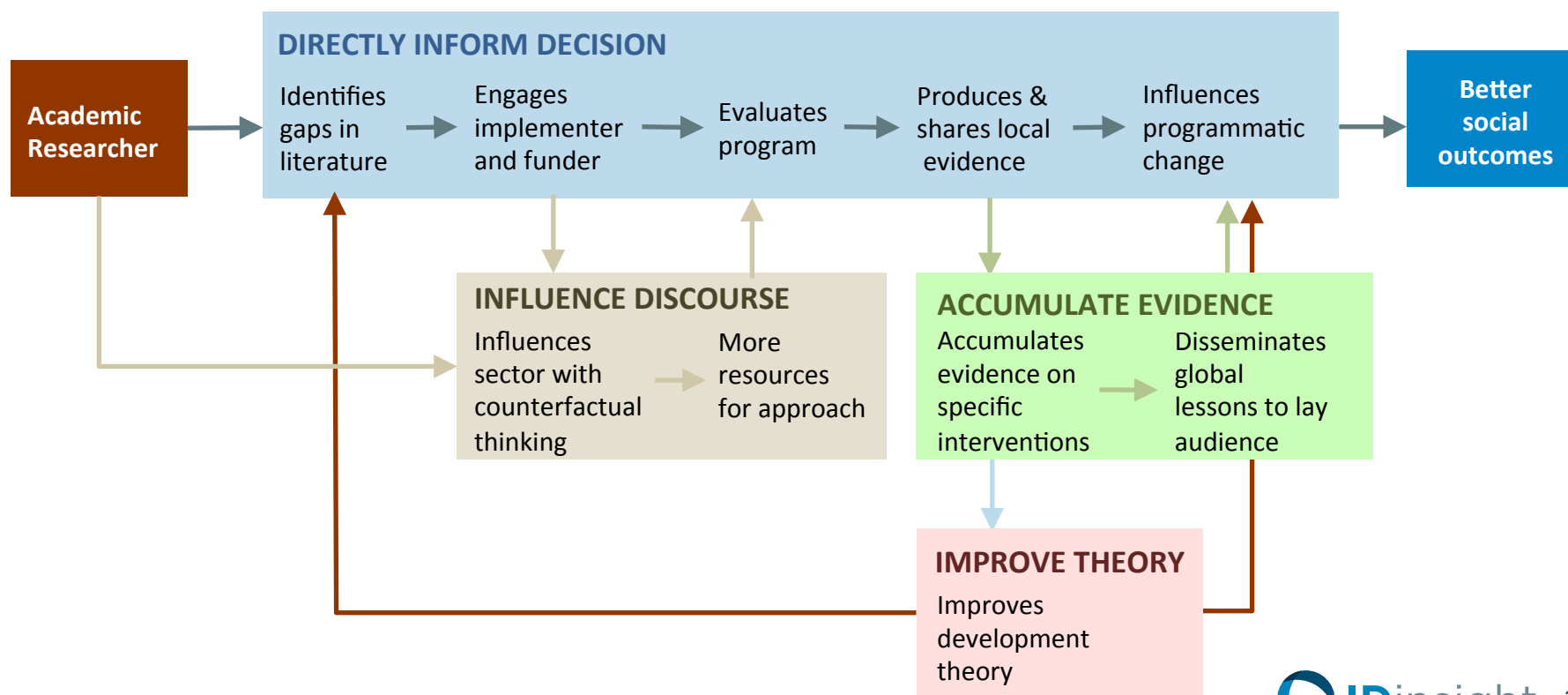
1. Directly inform implementing partner's decision
2. Influence development discourse
3. Accumulate global evidence



Knowledge-focused evaluation theory of change

Channels

1. Directly inform implementing partner's decision
2. Influence development discourse
3. Accumulate global evidence
4. Contribute to development theory



Knowledge-focused evaluation theory of change: Assumptions

Channel	Assumption	Assumption strength
Channel 1: Direct channel	Researcher incentives are aligned with the decision-making needs and constraints of implementers	Weak

“It must be acknowledged that the set of research questions that are most relevant to development policy overlap only partially with the set of questions that are seen to be in vogue by the editors of the professional journals at any given time...academic concerns that overlap imperfectly with the issues that matter to development practitioners.”

- Martin Ravallion, "Evaluation in the Practice of Development"

“The academic question was, ‘Does [the program] work?’ We know now that it does. But follow-on questions like, ‘What’s the optimal ratio of students to teaching assistants—should it be four to one; eight to one; twelve to one?’ That’s not a question you are going to get many academics excited about spending a year and a half and a couple hundred thousand dollars figuring out. But if you're a large organization running remedial education, you really ought to know the answer... we're doing less of that type of research, things that are distinctly non-academic but that are necessary for policy.”

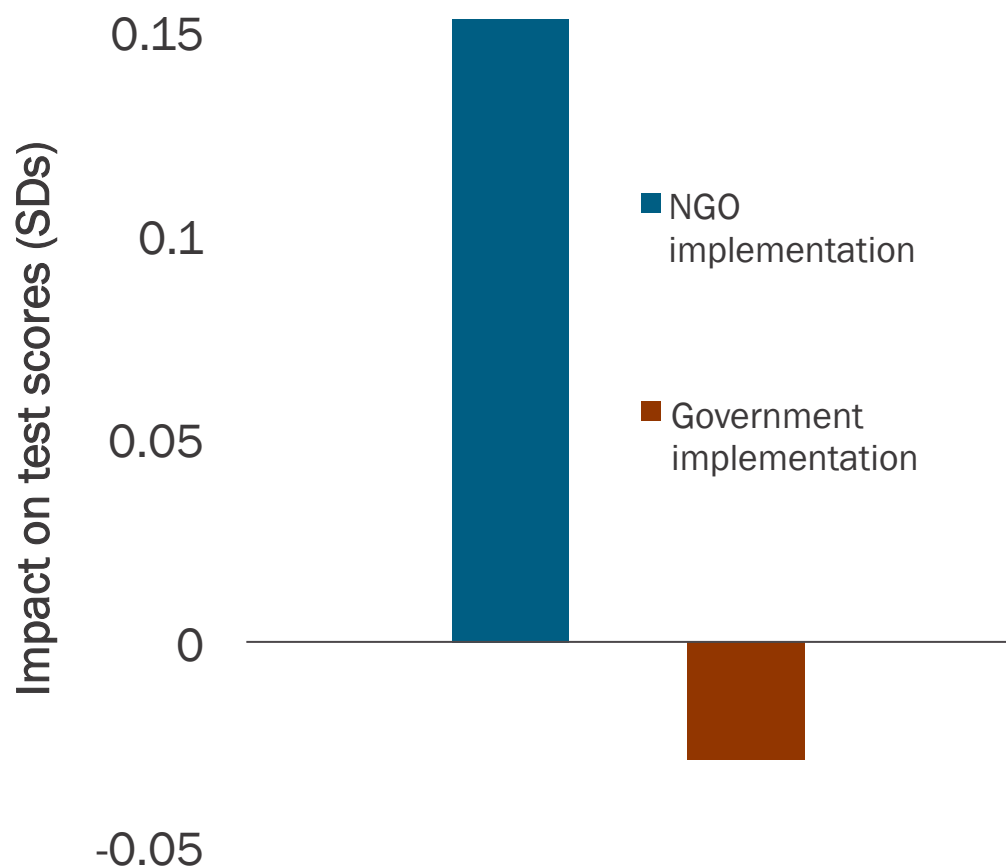
- Dean Karlan, on an evaluation of a remedial education program

Knowledge-focused evaluation theory of change: Assumptions

Channel	Assumption	Assumption strength
Channel 1: Direct channel	Researcher incentives are aligned with the decision-making needs and constraints of implementers	Weak
Channel 3: Accumulate global evidence	Findings from one impact evaluation can directly inform program design outside of that context	Weak

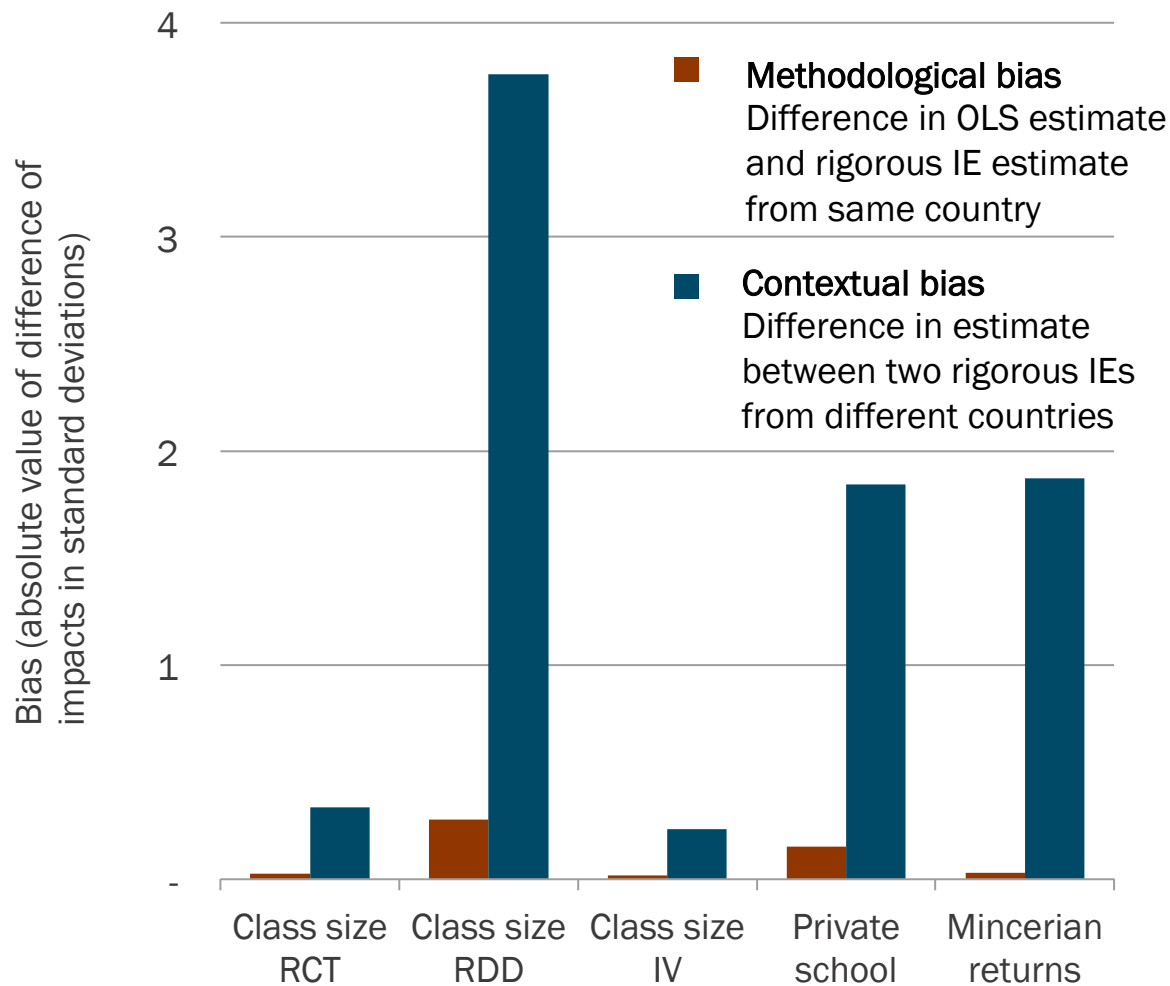
“Context” matters – Bold et. al. (2013)

NGO and Government implementation of the same contract teacher program in government schools in Kenya produced extremely different results



Context	Status
Intervention	Same
Geography	Same
Time	Same
School type	Same
Baseline characteristics	Same
Program implementer	Different (Govt vs. NGO)

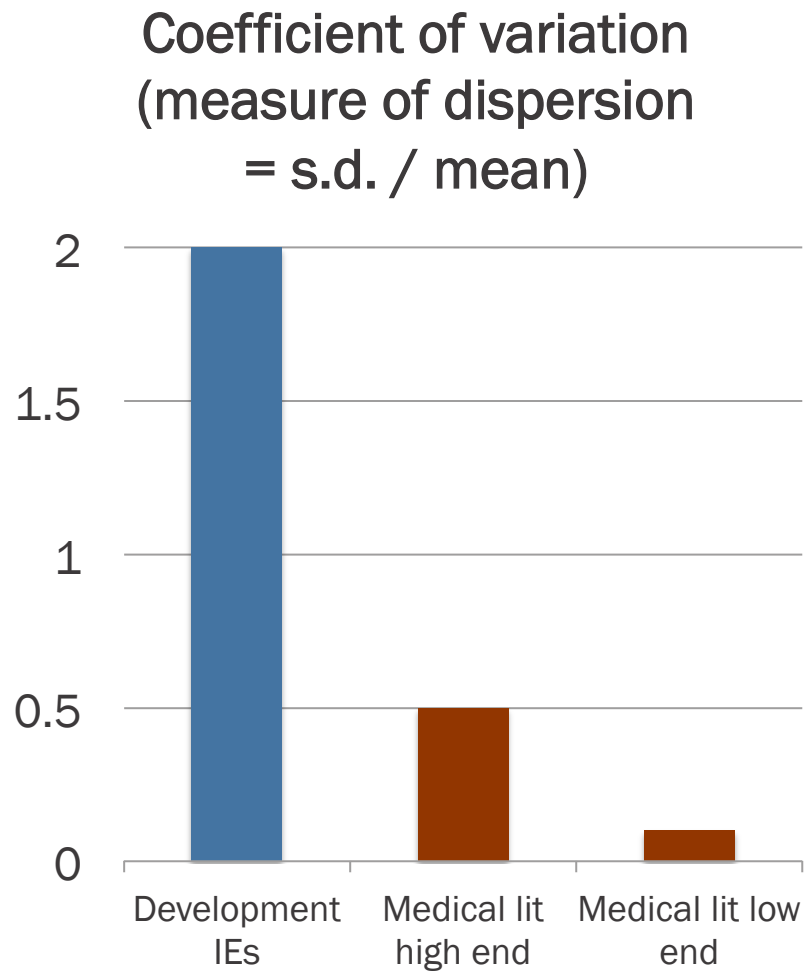
Contextual bias can be larger than methodological bias – Pritchett and Sandefur (2013)



Knowledge-focused evaluation theory of change: Assumptions

Channel	Assumption	Assumption strength
Channel 1: Direct channel	Researcher incentives are aligned with the decision-making needs and constraints of implementers	Weak
Channel 3: Accumulate global evidence	Findings from one impact evaluation can directly inform program design outside of that context	Weak
Channel 3: Accumulate global evidence	Replicating studies in many contexts can provide clear guidance on whether certain interventions “work”	TBD, but early evidence is weak

Large variation in intervention-outcome results across contexts – Vivalt (2015)



Predictive power of meta analysis to individual intervention is very low:

R²:

- 0.03 government implementer
- 0.07 NGO implementer

Appropriate role of knowledge-focused evaluations

- Continue to spend development resources on KFEs
- But focus those resources where they will be most useful
 - Questions about development theory
 - Instances with high external validity
- Not for: Informing the specific decisions of specific organizations in specific contexts.

Agenda

- About IDinsight
- Main Argument
- Status Quo: Knowledge-Focused Evaluations
- **New Paradigm: Decision-Focused Evaluations**
- Appropriate Roles
- Realizing the New Paradigm

Decision Focused Evaluations

Definition: an evaluation designed to:

- Inform a specific policy/programmatic **decision**
- For a specific **implementer**
- At a specific point of **time**
- In a specific **geography**
- For a specific **target population**

Comparing knowledge- and decision-focused evaluations

KFE approach

Researcher-driven

Longer timelines

Larger budgets

External

Maximizes internal validity



DFE approach

Demand-driven

Decision-making timelines

Budget in line with expected SROI

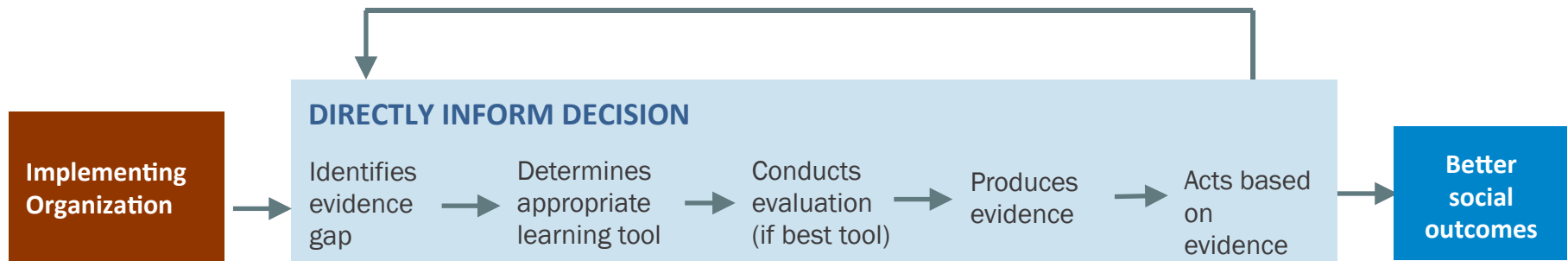
Embedded

Maximizes internally valid
(within decision-making
constraints and priorities)

DFE Theory of Change

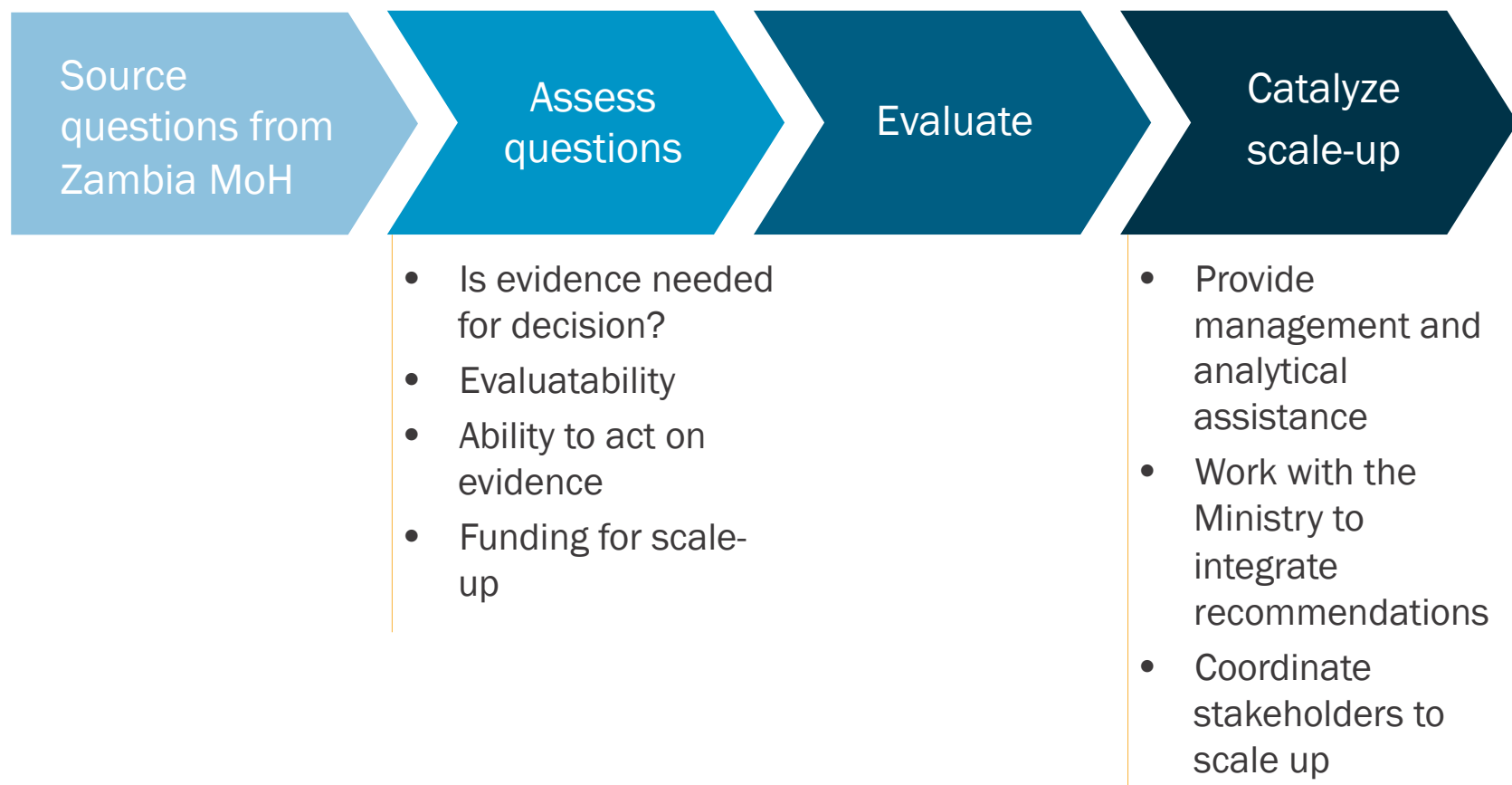
Channel

1. Directly inform implementing organization's decision
2. Influence development discourse
3. Accumulate global evidence
4. Contribute to development theory



Examples of DFEs by IDinsight

Example: Demand-driven evaluations for decisions





Demand-driven

Mama kits in Zambia: Demand-driven evaluation to inform national policy

Mama kits evaluation

Question: Can non-monetary incentives (“mama kits”) increase institutional deliveries?

Methodology: Cluster Randomized Evaluation

Sample size: 2,159 women at 30 health facilities

Duration: 4 months fieldwork

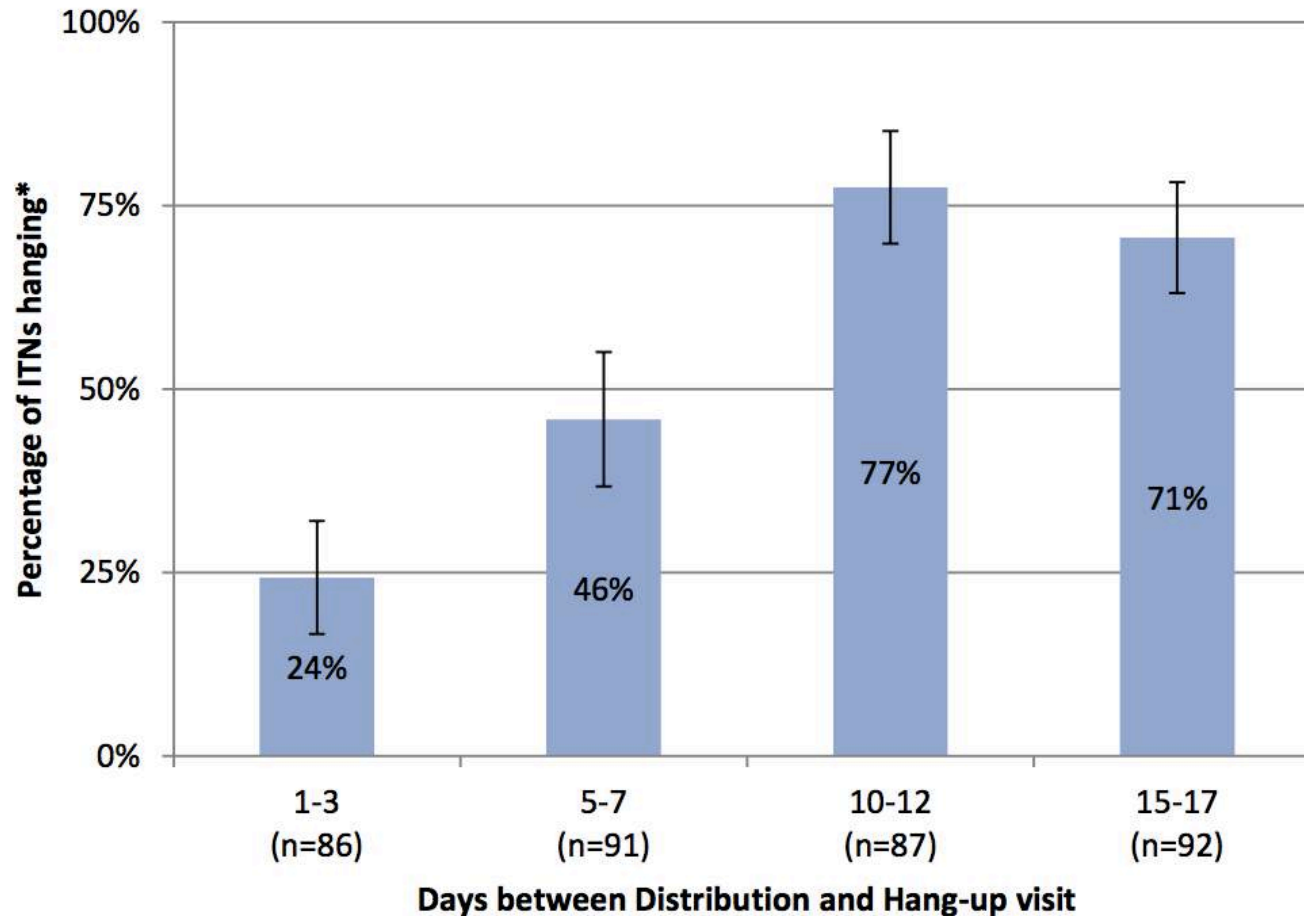
Results: Increased institutional deliveries by 47.4% Estimated
cost-effectiveness of \$3,414 per death averted

Action taken: MoH added “mama kits” to Zambia’s Essential Medicines List (9 months after evaluation start)

Insecticide-treated net (ITN) distribution evaluation

- Question:**
- 1) Can a community fixed point distribution approach quickly and cheaply achieve high ITN ownership and use?
 - 2) Do CHW follow-up "hang-up" visits increase ITN use?
- Methodology:** 5 Arm RCT (fixed point distribution vs. fixed-point distribution + CHW follow-up after 1-3 days; 5-7 days; 10-12 days; or 15-17 days)
- Sample size:** 593 HH
- Duration:** 3 months field work
- Result:**
- 1) High HH attendance (96%) and ITN retention (90%)
 - 2) CHW follow-up modestly increased short-term ITN use (13%)
 - 3) Fixed-point distribution with CHW follow-up reduce time and HR costs by ~25% versus door-to-door distribution (modeled)
- Action taken:** Zambia's ITN distribution guidelines adjusted to allow community point distributions (4 months after evaluation start).

ITNs: Optimizing CHW home visit delay



* Bars indicate 95% confidence intervals



Tailored

Evaluating d.light: Methodological flexibility improves evidence for decision-making

d.light evaluation - context

- d.light sells solar home systems in developing countries
- Received USAID DIV Stage II grant - required a rigorous evaluation of d.light's impact on health, and educational outcomes of d.light customers
- Objective of impact evaluation - inform USAID DIV decision of whether to provide Stage III "scale-up" funding for d.light

Evaluation design process

	KFE approach	DFE approach
Methodology	Randomized controlled trial ↓	Prospective matching + Dif-in-Dif. ↑
Business model	Offer products heavily subsidized / free ↓	Sell products at market prices ↑
Distribution partner	New partner w/no marketing experience ↓	Planned at-scale distribution partner ↑
Target customer	Poorer households than typical d.light customers in area where d.light products not sold ↓	Households that choose to purchase the d.light unit in existing d.light area of operation ↑
Question answered	Can d.light units improve welfare of selected HHs?	Does d.light's business model improve customers' welfare?

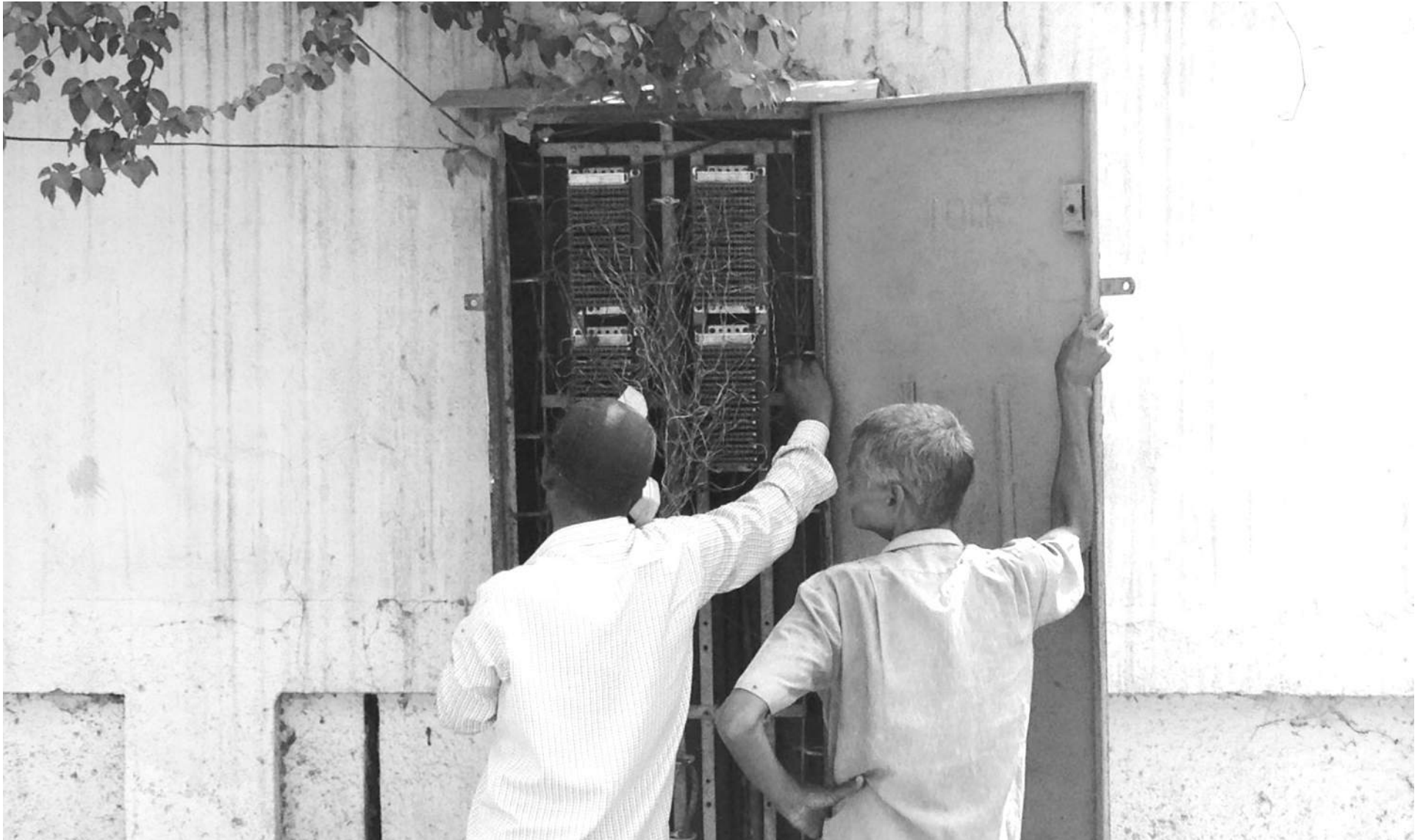
Solar home system evaluation

Question: Impact of d.light system on social outcomes

Methodology: Prospective matching with difference-in-differences analysis

Sample size: 1,000 HH

Duration: 9 months



Embedded

Concurrent Evaluation Cell (CEC) within Department of Energy, Government of Bihar, India

Concurrent evaluation cell – Bihar Department of Energy

Three key aspects:

1. Teams situated in government office; report directly to head of Department of Energy
2. Use a wide range of analytical tools, including decision-focused evaluations
3. Long engagement, shorter evidence cycles, more rounds of design, test, acting, repeat

CEC evidence directly leads to government action

Evaluation question: Do commissions for revenue collectors (RRFs) improve revenue collection in Department of Energy?

Results of RCT

1. RRFs increase revenue collection



Government action

Planning scale up RRFs in rural areas

2. RRFs do *not* increase meter reading



Develop new intervention (and evaluation) to incentivize RRFs

3. RRFs increase % of paying customers modestly

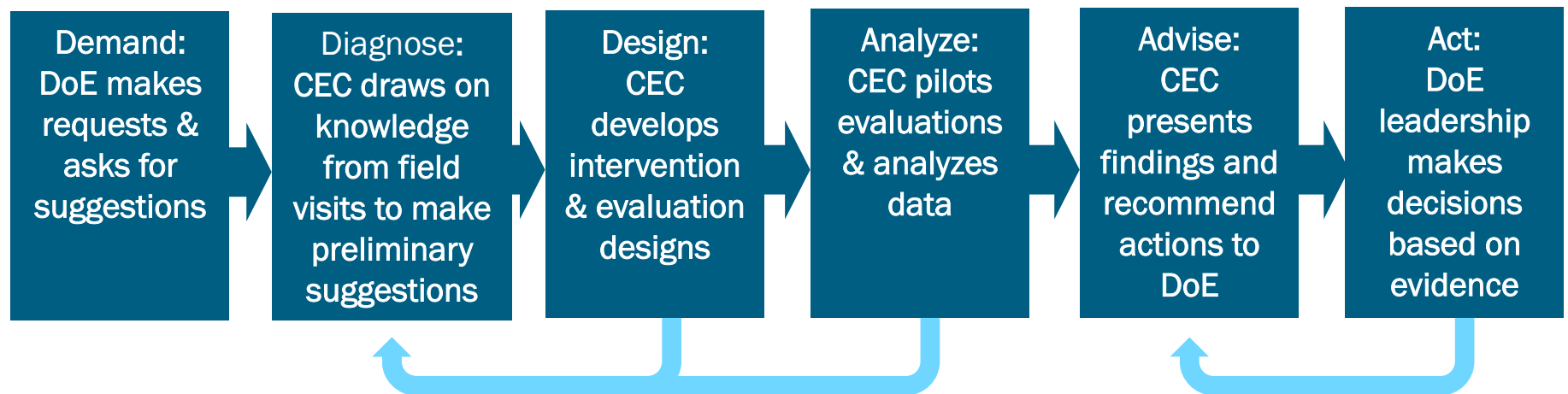


Develop new intervention (and evaluation) to incentivize RRFs to target non-payers

Government follow-up made easier, given embedded nature of CEC

Embedded CEC enables continuous experimentation

- CEC structure allows for **more experimentation and feedback loops** within a given budget and timeline
- This enables the CEC to:
 - Rapidly test multiple ideas
 - Quickly deploy appropriate monitoring or evaluation tool
 - Continuously refine ideas



Conditions required for DFEs to drive program improvement

- Being demand-driven and responsive are often not enough to drive evidence-informed decisions
- Acting on evidence depends on:
 - Client's authority to make policy decision on findings
 - Scale-up plan developed in advance - human and financial resources, stakeholder buy-in, and operational plan

Agenda

- About IDinsight
- Main Argument
- Status Quo: Knowledge-Focused Evaluations
- New Paradigm: Decision-Focused Evaluations
- **Appropriate Roles**
- Realizing the New Paradigm

Appropriate role of KFEs and DFEs

- All evaluations have a degree of both
- The future of impact evaluation must be characterized by
 - clear objectives
 - appropriate evaluation type

Primary objective	Evaluation type
Contribute to development theory	KFE
Explore the impact and mechanisms of very new intervention (e.g., large unconditional cash transfers)	KFE
Inform funder's decision to fund a specific organization's intervention	DFE
Inform implementer's decision of which programmatic variant to scale up	DFE

Combining KFEs and DFEs

- DFEs can build on KFEs
 - Review existing literature as starting point, source for ideas
 - Once KFE establishes link between proximate and end outcomes in a particular context, DFEs can stop at proximate outcome
- DFEs sometimes useful for informing theory and global knowledge

How to make this new approach a reality

This vision of the future of impact evaluation has two key components:

1. Appropriate use of KFEs in those circumstances that they are designed to add most value
2. Increase in DFEs and embedding of DFEs and other M&E tools within implementing organization's decision-making processes

Requires many changes in impact evaluation ecosystem

Agenda

- About IDinsight
- Main Argument
- Status Quo: Knowledge-Focused Evaluations
- New Paradigm: Decision-Focused Evaluations
- Appropriate Roles
- Realizing the New Paradigm

To realize this vision,
we need to
create a market
for decision-focused
evaluations

Building Supply

Market gap

1. Need for professional (non-academic) impact evaluators

2. Dearth of quality impact evaluation organizations operating on a demand-driven basis

3. Low capacity to generate and use evidence among implementers

Potential Solution

Fund professional tertiary education programs to train evaluators

1. Subsidize start-up funds for decision-focused evaluation orgs
2. Fast external quality review (e.g., 3ie) of eval design and PAP

Fund “build, operate, transfer” evaluation cells

Spurring Demand

Market gap

Solution

1. Implementers not incentivized to achieve impact



Tie scale-up funding to demonstration of impact over long time horizon

2. Evaluators not incentivized to prioritize decision-relevance



Publish the cost, length, and actions influenced by IEs in evaluation registries

3. Implementers lack funding to learn through decision-focused evaluations



Allocate dedicated portion of M&E/project funds for evaluations

Summary

IEs can have much greater impact in the future if:

1. Appropriate role of KFEs to better understand theory and human behavior
2. KFEs to examine new, potentially impactful development ideas
3. Dramatic growth of DFEs to directly inform decision-making of implementers
4. Closely embed DFEs and other learning tools within the decision-making structures of implementing organizations