

# USING THEORY-BASED EVALUATION

The case of evaluations in education

Stephen Taylor

DBE: Research, Monitoring & Evaluation

NES Seminar

28 June 2018



# Overview

- Definitions
- Benefits to using TBE/ToC in evaluating education programmes
- White's (2009) six principles for successful application of TBE
  - Illustrated using 3 recent education evaluations
- Lessons learned / Ways to improve future evaluations

# Definitions

- Theory-based evaluation
  - “The root idea of TBE is that the beliefs and assumptions underlying an intervention can be expressed in terms of a phased sequence of causes and effects (“program theory”)
    - Weiss (1997: 501)
  - “examining the assumptions underlying the causal chain from inputs to outcomes and impact”
    - White (2009: 3)

# Definitions

- A “Theory of Change” or “causal chain” maps out “how the intervention is expected to have its intended impact”. (White, 2009)

# Definitions: ToC

- Identifying long-term goals and the assumptions behind them
- Backwards mapping from the long-term goal by working out the preconditions or requirements necessary to achieve that goal--and explaining why.
- Voicing your assumptions about what exists in the system without which your theory won't work, and articulating your rationales for why outcomes are necessary preconditions to other outcomes.
- Weighing and choosing the most strategic interventions to bring about your desired change.
- Developing indicators to measure progress on your desired outcomes and assess the performance of your initiative.
- Quality review should answer three basic questions: Is your theory 1) plausible, 2) “doable” (or feasible), and 3) testable?
- Writing a narrative to explain the summary logic of your initiative.

Clark & Taplin (2012)

# Our experience in DBE

Evaluation	Year completed	Type of evaluation	Key lessons
ECD Diagnostic Review	2012	Diagnostic	Further program design work needed
Grade R programme	2013	Impact	Negligible effect of Grade R on later learning, except in more affluent schools and better functioning provinces
National Schools Nutrition Programme	2016	Implementation	Specific ways to strengthen programme implementation to increase likelihood of impact
Funza Lushaka Bursary Programme	2016	Implementation	Specific ways to strengthen programme implementation to increase likelihood of impact
CAPS (curriculum)	2017	Implementation	Specific ways to strengthen programme implementation to increase likelihood of impact
Early Grade Reading Study	2018	Impact (with mixed methods)	Daily lesson plans, African language reading materials plus on-site coaching yields largest impact
Learner Transport Programme	Ongoing	Implementation	

# Our experience in DBE: Benefits of TBE/ToC

- To inform programme design
  - E.g. identify weak links (great meal, eaten too late)
  - Often retrospective, but can inform improvements to programme design
  - Spillover benefits for thinking about design



# Benefits of TBE/ToC: Survey Instrument Development

A	B	C	D	E
		Indicators	Specific indicators	Modality
<b>Final outcome</b>	<b>Final Outcome</b>	Reading Ability	non-word decoding	EGRA
		Reading Ability	Word Recognition	EGRA
		Reading Ability	Vocab	Oral pictorial
		Reading Ability	Fluency	(count number of correct words per minute)
		Reading Ability	Comprehension (?)	
		Reading Ability	ANA results	ANA
<b>Intermediate Outcomes</b>	<b>Teaching behaviour</b>	Changed teaching practice		Classroom observation
		Teacher motivation	use of workbooks	document review
		Teacher motivation	preparation of lesson plans and curriculum documents	document review
		Teacher motivation	availability of CAPS documents	document review
		Teachers knowledge	pedagogy: e.g. test decoding skills	short questions within teacher survey
	<b>Parent beliefs/behaviour</b>	Teachers knowledge	reading fluency (??)	e.g. Schonell's reading age (??)
		Parent behaviour	check homework	Survey pupil? // workbook review
		Parent behaviour	practice reading	pupil survey
		Parent behaviour	number and content of parent meetings	principal questionnaire and document review
		Parent belief - school can improve		parents focus group discussion: 2 SGB and >2 I
<b>Characteristics</b>	<b>Demographics</b>	Parent belief - responsibility to act		parents focus group discussion: 2 SGB and >2 I
		Parent belief - value to learning/reading		parents focus group discussion: 2 SGB and >2 I
		Home language		Pupil/teacher survey
		Gender		Pupil/teacher survey
		Age		Pupil/teacher survey



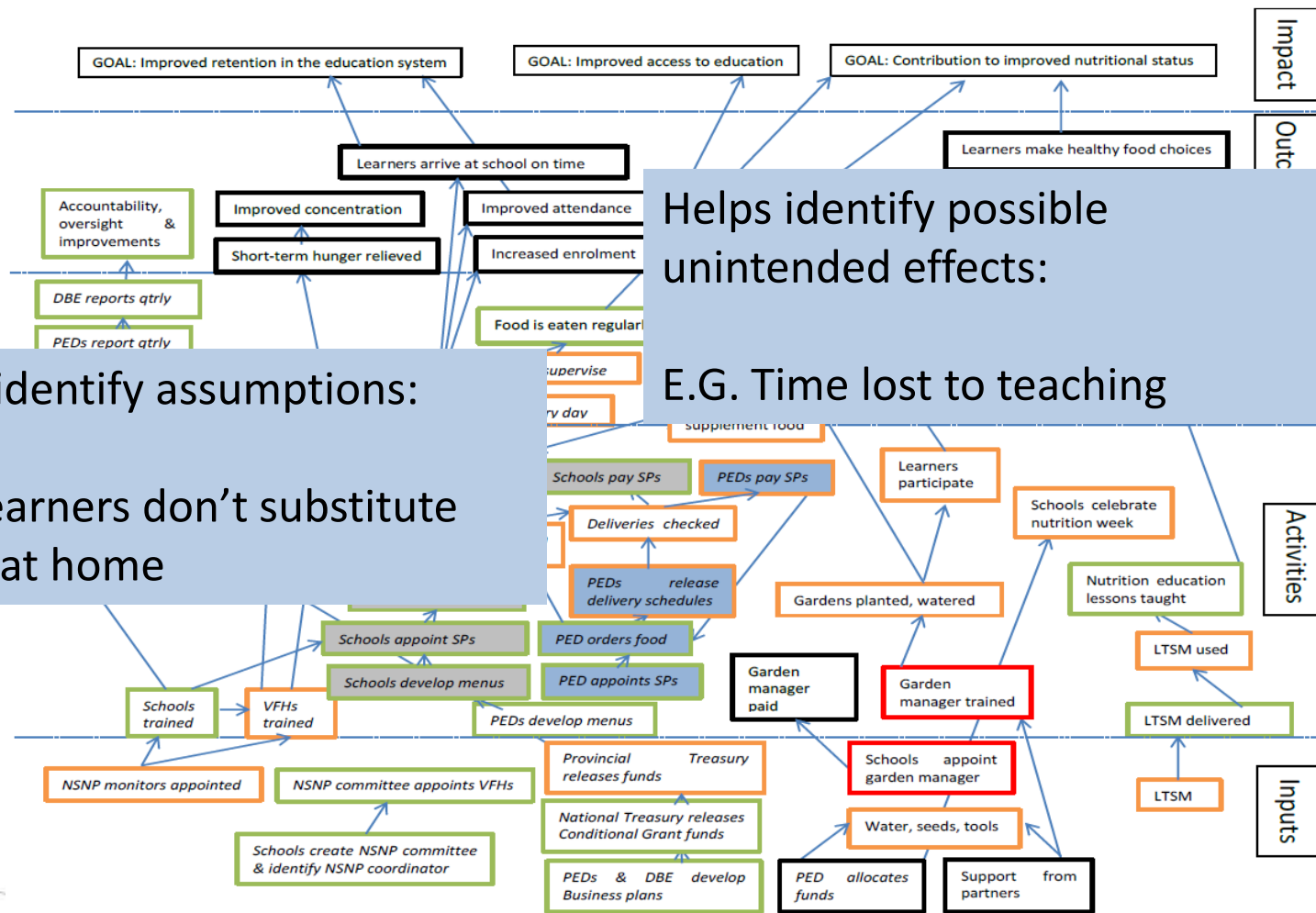
# Benefits of TBE/ToC: Guiding the Analysis Plan

T1: TRAINING	T2: COACHING	T3: PARENTS
<b>THREATS TO VALIDITY</b>		
<ul style="list-style-type: none"> <li>- Balance tests</li> <li>- Attrition</li> </ul>		
<b>FIDELITY OF IMPLEMENTATION</b>		
Attendance at training	Number of coaching visits	Parent attendance
<b>INTERMEDIATE OUTCOMES</b>		
Do teachers feel supported? Changes in classroom practice	Do teachers feel supported? Changes in classroom practice	Parent attendance at meetings Changes in attitudes and practices
<b>IMPACT ON LEARNING OUTCOMES</b>		
Positive or negative impacts on HL literacy (Primary objective), English and Mathematics		
<b>IMPACT ON SUB-GROUPS</b>		
<p><b>Does impact depend on:</b></p> <ul style="list-style-type: none"> <li>Baseline learner proficiency</li> <li>Learner gender</li> <li>Teacher characteristics</li> <li>School types (location, size, socio-economic status)</li> <li>Parent/home characteristics</li> </ul>		
<b>ROBUSTNESS CHECKS</b>		
<b>COST-EFFECTIVENESS ANALYSIS</b>		

# White's (2009) six principles for successful application of TBE

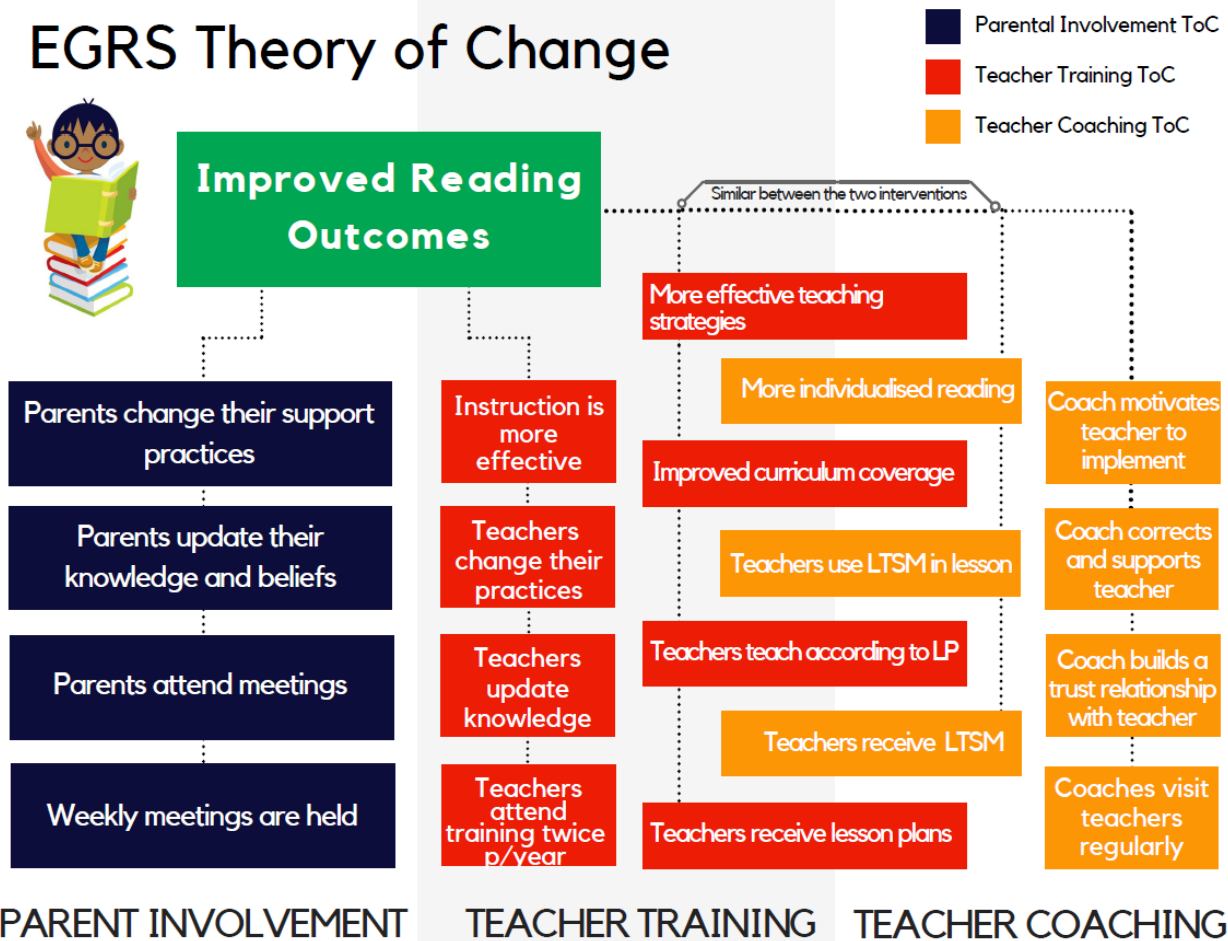
- 1. Map out the causal chain (programme theory)
- 2. Understand context
- 3. Anticipate heterogeneity
- 4. Rigorous evaluation of impact using a credible counterfactual
- 5. Rigorous factual analysis
- 6. Use mixed methods

# 1. Map out the causal chain (programme theory)



# 1. Map out the causal chain (programme theory)

## EGRS Theory of Change



# 1. Map out the causal chain (programme theory)

- Limitations
  - Useful tool, but not definitive
  - May need to be revised after the evaluation
  - Level of complexity needs to be decided

## 2. Understand context

- How realistic are your assumptions implicit in the ToC?
- In EGRS we looked at:
  - Initial reading outcomes of children
  - Parent educational attainment
  - Frequency of support from district curriculum advisors
  - etc

# 3. Anticipate heterogeneity

- Pre-Analysis Plan

## PUPIL LEVEL

- |  |                       |
|--|-----------------------|
| 1. Pupil gender  | [pupil baseline test] |
| 2. Pupil age   | [pupil baseline test] |
| 3. Pupil baseline performance                                  | [pupil baseline test] |
| 4. The interaction between baseline performance and class size | [pupil test and SPQ]  |

## TEACHER LEVEL

- |   |                        |
|---|------------------------|
| 1. Baseline teacher characteristics:                              |                        |
| a. Education level  | [TRQ – 3(a)]           |
| b. Performance in the teacher reading test                        | [Teacher written test] |
| c. Years of experience  | [TRQ – 3(c) and 3(e)]  |
| d. Own background reading   | [TRQ – 4]              |
| 2. Baseline teacher presence                                      |                        |
| 3. Whether teacher spent time after class hours preparing lessons |                        |
| 4. Previous training opportunities                                | [TRQ – 6(a) and 6(b)]  |



# 3. Anticipate heterogeneity

- Pre-Analysis Plan

## SCHOOL-LEVEL

- |  |                                    |
|--|------------------------------------|
| 1. Baseline school average reading performance   | [administrative data]              |
| 2. Socio-economic background of school           | [administrative data]              |
| 3. School resources                              | [SPQ: 4(a), 6, and 7]              |
| 4. School location                               | [SPQ – 4(b)]                       |
| 5. Pupil-teacher ratio                           | [SPQ 2(c) and administrative data] |
| 6. Overall educational background of the parents | [SPQ – 3(f)]                       |
| 7. Overall employment of parents                 | [SPQ – 3(g)]                       |

## PARENT LEVEL

- |  |                                    |
|--|------------------------------------|
| 1. Identity of the parent/guardian:<br>E.g. single mother/grandmother; child-headed household. | [PRQ - 2]                          |
| 2. Education level   | [PRQ – 4]                          |
| 3. Baseline written ability of parent  | [PRQ – whether they can enter Q11] |
| 4. Baseline involvement in child's learning  | [PRQ – 5,6 and 9; 12 and 13]       |

# 4. Rigorous evaluation of impact using a credible counterfactual

- Not only RCTs!
  - “Gold standard” is quasi-experimental using administrative data
- RCTs and other quasi-experimental studies can be set up to rigorously explore mechanisms (a limited number though)
- Ideal is plausible counterfactual (impact) and detailed how and why analysis (using mixed methods)

# 4. Rigorous evaluation of impact using a credible counterfactual

- Grade R impact evaluation
  - Impact only (with NB subgroup effects)
  - Limited information for how to improve
- NSNP/FLBP/CAPS
  - Good analysis of pre-conditions for likely impact
  - Good recommendations for how to improve programme implementation
  - We don't know impact
- EGRS
  - 3 levels of analysis (macro, mezzo, micro)

# 5. Rigorous factual analysis

- What percentage of schools served their meals on time?
- What percentage of parents attended parent meetings?
- Do teachers feel supported?
  - Counterfactual also useful here...
- Are provided materials visible in classrooms?
  - Counterfactual also useful here...
- How many times did coaches visit schools?
  - In urban versus rural areas?

# 6. Use mixed methods

- Wide range of methods
- In education the ultimate “black box” is the classroom
  - Difficult to observe on large scale
  - But limitations of small scale (e.g. class size finding)
- Qualitative and quantitative methods should interpret each other
  - Generating and testing hypotheses

# Final thoughts

- Doing good TBE relies on capacity
- Budgets are a constraint
  - CAPS evaluation
- Better administrative data can really help
- TBE is difficult and cannot solve all problems but adds a lot of value
  - Weiss: We should try to overcome the obstacles
  - Same applies to experimental IEs