

Do higher pay and better training improve teaching? Evidence from Indonesia

SIEF Regional Impact Evaluation Workshop

Manila, December 2, 2008

Halsey Rogers (World Bank)

With Amanda Beatty (World Bank), Benny Karyadi (Indonesian Ministry of National Education), Karthik Muralidharan (UC San Diego), and Menno Pradhan (Amsterdam Institute for International Development)

Overview

- Context for evaluation
- Evaluation design and goals
- Benefits of these evaluations
- Alternate designs and challenges

Context for evaluations: Indonesian teacher reforms

- Major set of reforms to upgrade teacher skills and professionalism in Indonesia
 - Launched by Teacher Law passed by Parliament in 2005
 - Ministry has been designing and implementing, with WB help
- Major features of the reform:
 - Requires certification of all 2.7 million teachers by 2016
 - Certification requires 4-year college training or equivalent, as well passing evaluation of teacher's portfolio
 - Certified teachers receive doubling of base salary
 - Reforms linked with other changes as well (e.g., much larger bonus for posting to remote areas)

Context for evaluations: Multifaceted M&E program

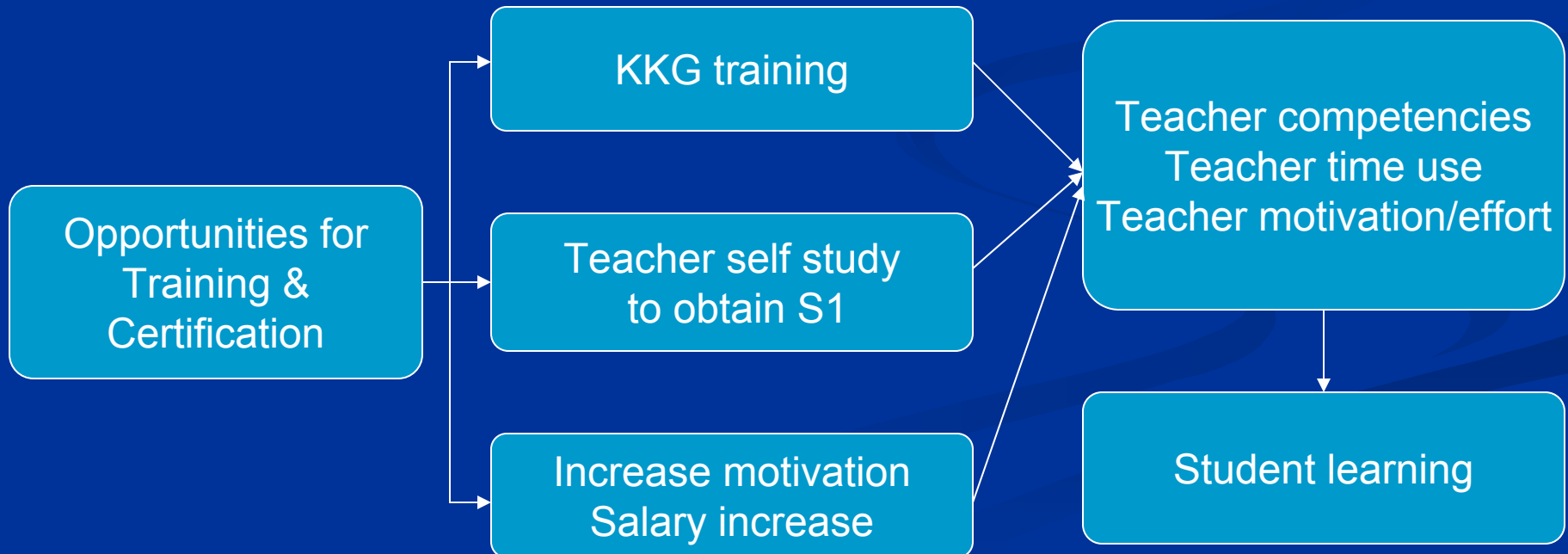
- World Bank coordinating with government on evaluations along various points of results chain, e.g.:
 - TIMSS video study of teacher practice in the classroom (*careful, rigorous observation*)
 - Rapid process evaluation of certification (*case study*)
 - Evaluation of changes in the quality of entrants into teaching profession (*difference-in-difference using observational data*)
 - Evaluation of remote-area bonuses (*diff-in-diff, observ. data*)
- For evaluation of other components, RCTs seemed to be most appropriate methodology
 - Experimental evaluations of effects of salaries & training

Parallel experimental evaluations

- Parallel RCT-based evaluations are planned to assess effects of two parts of the reforms:
 - **Certification (salary) experiment:** randomize accelerated access to the certification queue & to possible salary increase
 - **BERMUTU training experiment:** randomize access to cluster-based teacher training financed under WB-supported BERMUTU project
- Together, they cover the two policies that are often recommended for improving teaching quality:
 - increase salaries
 - train teachers better

Tracing the pathway of (potential) impact

- There are three main channels through which teacher certification and Bermutu are expected to improve outcomes
 - Increased motivation and effort of teachers due to higher salaries
 - Professional upgrading of teachers without 4-year college degrees
 - Continuous professional development through enhanced teacher cluster groups – facilitated through grants to groups
- **Critical question – how much does this actually take place?**



Goals of evaluations

- Determine the impact of certification/salary increase and cluster-based teacher training on:
 - Teacher motivation and effort
 - Teacher competency and skills
 - Student learning outcomes
- Understand why and when there was an impact
 - Through which channel is impact obtained?
 - How does the behavior of key actors change?
 - In which circumstances is program more effective?
 - How does program impact vary with baseline characteristics?

Evaluation design – Critical elements

- **Randomized allocation of programs** to treatment and control groups (in 20 evaluation districts)
 - Ensures that schools being studied are identical in all ways except for the introduction of the program
 - The treatment will provide for **accelerated** access to certification and training programs (but programs are not modified in any way)
 - Control group will proceed according to “business as usual” (NOT excluded from normal program)
- **Tracking of process and outcome variables** before and after introduction of program to link investment in teachers to learning improvements

Design of certification/ salary-increase experiment

- Select 10 nationally representative districts (out of 451) as experimental districts
- Randomly select 600 junior secondary schools for tracking of teacher knowledge/behavior and student learning outcomes
- **Treatment** schools (200): teachers gain immediate access to certification process, if ready → salary increase sooner
- **Control** schools (400): teachers wait in usual queue (queuing by seniority) for access to certification process
- **Benefit:** Allows us to assess effect of salary increase without confounding factors (like seniority)

Design of training experiment

- Select 10 districts out of 75 in BERMUTU project
- Randomly select 600 primary & junior secondary schools for tracking of teacher knowledge/behavior and student learning outcomes
- Randomly select teacher clusters for grant support out of those submitting qualifying applications
- **Treatment** schools: Those whose teachers are in clusters receiving grants in year 1
- **Control** schools: Those whose teachers are in clusters not receiving grants until later (or not at all)
- **Benefit:** allows evaluation of training without confounding factors (e.g., level of activity of teacher cluster)

Outcome variables

- In both control and treatment schools, plan to regularly measure outcomes
- Final outcomes:
 - Student learning achievement
- Intermediate outcomes:
 - Teacher competency (measured by subject knowledge)
 - Teacher motivation (measured through interviews and observed behavior, such as attendance)
 - Teacher time allocation

Benefits of evaluation

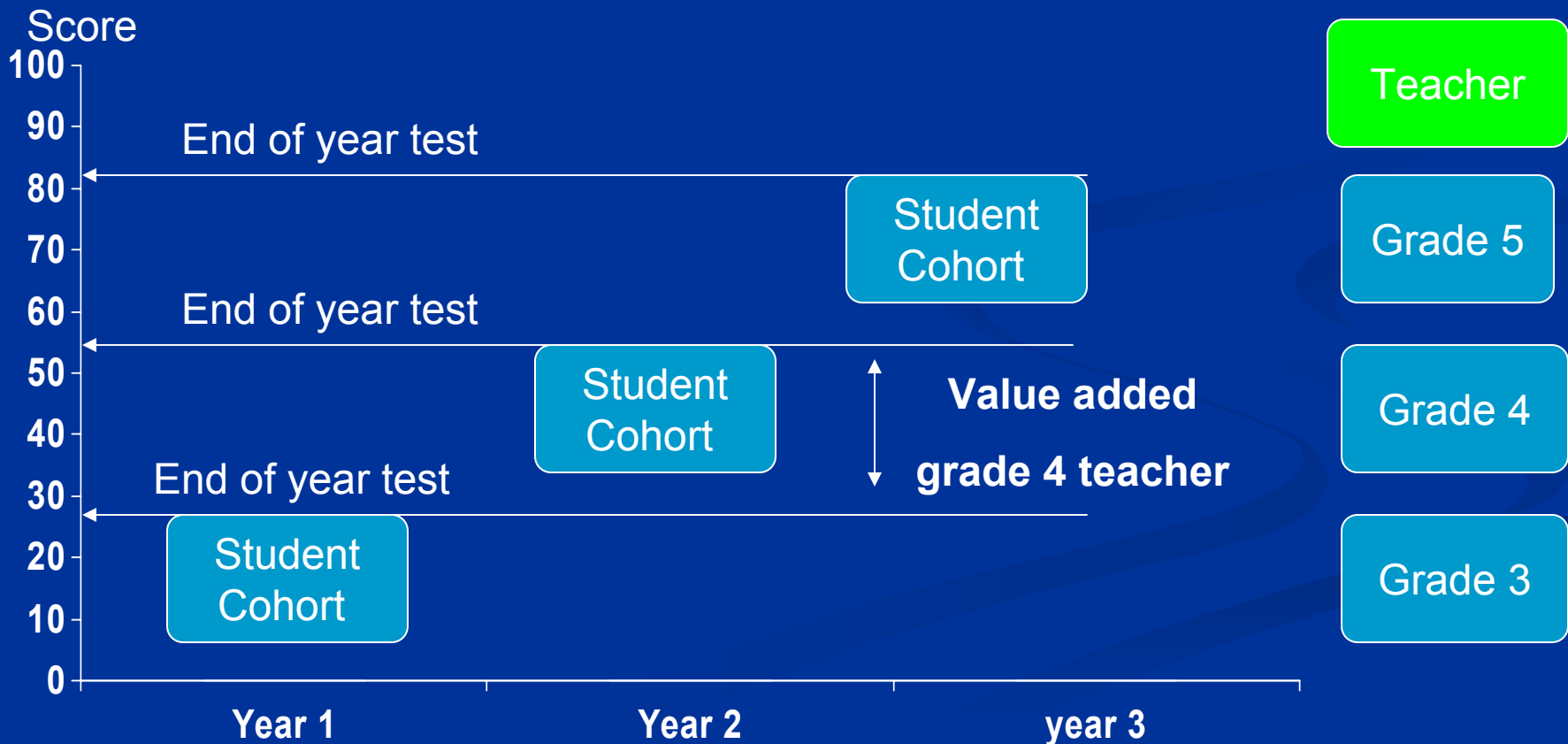
- Immediate benefits
 - Clarify thinking about the precise channels of impact
 - Responds to demand for impact evaluation of education policy
- After baseline (6 – 12 months from now)
 - A snapshot of learning achievements and teacher competencies and factors associated with better outcomes (20 districts)
- 1-2 years after baseline:
 - Process evaluation of Bermutu training and certification
 - Impact of certification/pay raise, and cluster training on teacher competencies, effort, and student learning
 - Understanding of differential program effectiveness, which can help to improve program in subsequent years
- 3+ years after baseline
 - Impact of certification program on non-college-grad teachers' competencies, effort, and student learning (through upgrading)
 - Longer term impact of certification/pay raise on college-educated teachers, and of cluster training on all teachers

Additional benefits of evaluation

- Systematic collection of **longitudinal data on schools, teachers, and students** in a representative sample of schools
 - Can match each student to school, teacher, classroom, and peer characteristics
 - Can measure the “value added” of each teacher and the relationship between teacher characteristics and student learning outcomes
 - Allows tracking of long-term effects
 - Will support strategies in the education sector much more broadly by building data-driven understanding
- Putting evaluation structure in place also provides an opportunity to **pilot alternate forms** of teacher incentives and evaluate these using the same measurement tools.
 - Included in component 3 under Bermutu project

Tests used in IEs measure value added by individual teachers

- Link learning gains to teacher(s) who taught the students
- Multi-year tracking and testing of students and teachers



Alternative designs that were considered and rejected

- Implement certification and training experiments in the same set of districts
 - Too risky, due to possibility of delays and complications of sample selection
- Implement certification/salary increase experiment at both the primary and junior secondary level
 - Not possible, due to shortage of primary teachers with 4-yr degree

Challenges faced

- Core elements have remained constant, but specific design features have needed to change due to:
 - Implementation schedule
 - Government preferences
- Conditions on the ground have made certain evaluation approaches untenable
 - E.g., RDD to evaluate certification or remote-areas bonus
- Coordination and communication have been challenging at times, despite strong counterparts
 - Very active sector, with different departments/ministries involved

Some remaining challenges

- The many demands on government reduce the time available to plan and manage these evaluations
 - Extent of reforms means counterparts can be stretched thin with implementation, leaving less time for evaluation
- Delays in implementation of evaluation can reduce viability of the evaluation design
 - Vanishing number of uncertified college-graduate teachers for experimental sample at some levels

Summary

- Exciting opportunity to:
 - Evaluate key portions of major government education program
 - Shed light on problems of broad international interest, with unique data and opportunities
 - Put in place a structure for longer-term measurement, experimentation, and improvement