

**Three Experiments to Improve
Learning Outcomes in Tanzania:**

**Delivering capitation grants better and
testing local cash on delivery**

**Final Draft
June 2012**



Three Experiments to Improve Learning Outcomes in Tanzania: Delivering capitation grants better and testing local cash on delivery

Final Draft
June 2012

1. Introduction

Overall student learning levels remain extremely low across East Africa, despite a decade plus of major reforms and significant new investments in public education. In Kenya, Tanzania and Uganda, recent nationwide surveys show that large majorities of children are unable to read or do arithmetic at the required level (Uwezo at Twaweza, 2011). Children cannot learn new skills without a foundation in basic literacy and numeracy; without this foundation they are also denied the opportunity to develop fully in the future. Across East Africa, students entering secondary and tertiary education institutions as well as labor markets are generally woefully ill prepared. While these challenges are well known, existing reforms and aid instruments have largely failed to improve the situation.

At present, the two main approaches used by governments to improve the quality of education in East Africa are to strengthen teacher training and to disburse a capitation grant for books and related activities to schools. On the former, several studies show formal levels of teacher qualification to be weakly correlated to performance. On the latter, it is difficult to establish the impact of capitation grants on improving learning outcomes because the full amounts of the grant have not consistently reached schools and because they have not been rigorously evaluated. In Tanzania, the amounts reaching schools are well below policy of USD 10 per primary pupil per year and are declining, from USD 6 in 2004 to USD 4 in 2009 to less than USD 2 in 2011 (Education PETS in Tanzania 2004, 2009; see also Reinnika and Svensson, 2004 for Uganda).

At the same time, even if funds were to flow well, under current arrangements no one is held accountable or incentivized to achieve learning. Administrators and teachers are paid regardless of their attendance or performance; quality assurance systems such as the inspectorate function poorly; and appointments of education administrators do not take learning outcomes into account. The lack of adequate attention to accountability and incentives may in part explain why increased budgets for education have not resulted in improved learning outcomes. So while government programs have largely focused on providing *educational inputs*, recent evidence suggests that it may be more effective to incentivize the delivery of *learning outcomes*, particularly at the local level (see Glewwe and Kremer 2005, and Kremer and Holla 2009, as well as the Annex to this note).

In short, we see three key challenges to improving education in East Africa:

- how to consistently get resources to the lowest (school/community) levels,
- how to effectively invest (government and donor) resources by gearing the education process to emphasize learning outcomes rather than educational inputs,
- how to generate rigorous evidence of what works and have it inform national education debate, policy and practice.

Twaweza seeks to address all three challenges. In this note, we propose to conduct two separate interventions and a combination of the two interventions to improve the channeling of resources to public primary schools for quality improvement and the structuring of incentives to improve learning outcomes. Our aim is to implement and rigorously test a set of interventions involving resources,

information and incentives, and to engage a community of researchers, think tanks, government, teachers union, civil society and donors in the process. Ultimately, we aim to inform national policy and stimulate public debate with evidence from interventions. As such, particular attention will be given to the involvement of members of parliament (MPs) and leadership of local councils, and to publicizing the findings.

This note is organized as follows: Section 2 outlines the Intervention design involving three experiments. Section 3 discusses policy considerations. Section 4 lists key partners. The Annex provides a literature review of related studies.

2. Intervention Design

To summarize, our main hypotheses are that:

- Increased information to citizens will improve disbursement and use of Capitation Grants (CG)
- Increased CG will improve student learning outcomes
- Local Cash on Delivery incentives will improve student learning outcomes
- COD will improve student learning outcomes more than CG
- Targeting funding at the lower structures, i.e. schools, teachers, and communities, will improve student learning outcomes more than higher structures, i.e. districts.

A randomized control trial (RCT) experimental design with random allocation of schools to treatment arms is envisioned to ensure methodological rigor in isolating causal drivers and attributing impact (Banerjee and Duflo, 2011), complemented by qualitative approaches that seek to understand why and how things work.

In total three experiments are planned to be undertaken in parallel. The first experiment will focus on making *existing policy* of delivering capitation grants (of USD 10 per pupil/per year) work better, by channeling funds in full and more effectively to public primary schools, and to test the effects of information provision about the grants. The evaluation will seek to measure the extent to which the funds reach schools, the level of citizen engagement on the use of funds, and ultimately the impact of funds and information on improving learning outcomes.

The second experiment will test the impact of paying for improved learning outcomes at the local level (at about USD 20 per pupil who does well on the Uwezo assessment), an *innovation* that has not been tried in education in East Africa below. Because the same measure of learning outcomes will be used in both experiments (the Uwezo Grade 2 level literacy and numeracy tool), the experiments will allow us, to some extent, to compare the existing policy of provision of input funds *upfront* with the innovation of paying for outcomes *upon delivery*.

The third experiment will involve a combination of the first two experiments i.e. provision of capitation grant inputs upfront plus paying for outcomes on delivery of outcomes.

We will not prescribe any specific activities to improve learning, other than providing access to information on methods used by others. What works to improve student learning may vary across districts, schools, teachers – 100 different teachers may try 100 different methods. Our experimental design seeks to answer not “which specific activity helps students learn the most?” but the impact of input provision and outcome based payment types of approaches at local levels, and the role of both information and citizen engagement in the process.

The preparatory work for the experiments, including piloting of materials and engagement, will take place throughout 2012, in consultation with key parties. The full field interventions are expected to begin in early 2013. Because the development of new approaches and their uptake, including necessary tweaks and adaptations, can take time, these experiments will be implemented over a minimum of two years, and in all likelihood over a longer period. This can also help ascertain the sustainability of certain interventions over time, and as they are scaled up.

Intervention 1 – Making Capitation Grants Flow (CGF):

According to policy, the government is to disburse USD 10 per pupil per year to school accounts, which the school leadership with guidance from the school governance committees may use for improving quality of education. However, in the past decade the funds sent to schools have been well below policy levels (and in fact are declining), disbursements have been unpredictable, and levels of information at district and community levels have been inadequate. The situation has been compounded by the fact that the capitation grant is funded out of different budget lines, and at times districts failing to release funds meant for schools.

Our CGF intervention seeks to:

- demonstrate a more effective way to transfer capitation grants (CG) such that they reach schools in full (USD 10 per pupil/per year) in a timely and predictable manner;
- test the effect of providing information on reach and use of funds, including citizen engagement with the funds; and
- test the impact of providing capitation grants and information on improving basic literacy and numeracy.

The experimental design will seek to assess each of these three aspects, and is expected to have the following four treatment arms. The control arm will consist of current practice; no effort will be made by us to alter government provision of funds and information in the treatment or control arms.

Capitation Grant Flow intervention

Disbursed to	Info level	Arm	Experiment Arm
<i>Indirect</i> to district (local government authority) education accounts	low	1	Indirect CG disbursement to district education account, with normal (low) public information, as per current policy
	high	2	<i>Indirect CG disbursement and information:</i> to district education account with additional public information to schools and communities
<i>Direct</i> to school accounts (with districts informed)	low	3	<i>Direct CG disbursement:</i> to school accounts, with normal (low) public information, as per current policy
	high	4	<i>Direct CG disbursement only:</i> to school accounts
		C	<i>Control:</i> current practice, including whatever funds and information are provided by government

The capitation grants will be provided by Twaweza in two disbursements per year (at USD 5 each per pupil) on set dates. Funds will be transferred by electronic transfer into already established bank accounts. Schools may only use the funds, consistent with present policy, for improving quality (i.e. not for salaries or major construction). Twaweza will not establish any special systems for the planning, use, accounting and reporting of these funds; schools are to use existing government policies and institutional arrangements, which provide a prominent role for school committees and

head teachers and oversight by the district education offices. By varying where the funds are disbursed, the experiment will test to see if there is a difference, at the school level, between the indirect transfer of funds to district education bank accounts (as per current policy) and the direct transfer to existing school accounts.

According to current policy, information about capitation grants is meant to be made public by posting it in national newspapers and on school noticeboards. In practice, adherence to this policy has been somewhat erratic. By varying the levels of information provided, the experiment will seek to test if there is a difference between fully implementing how the government would normally provide information and a more elaborate, popular communications initiative focused on teachers, parents, pupils and the community. In the latter case, the type of information provided will reflect particular hypotheses in Twaweza's Theory of Change (see www.twaweza.org). The aim is to assess the extent to which information enhances citizen engagement and accountability, and whether these contribute to better use of funds and improved learning. Because some aspects of the intervention are likely to be reported in national media, the research design will need to find ways to account for potential contamination or spillover effects.

Intervention 2 – Local Cash on Delivery:

Twaweza will modify the Cash on Delivery (COD) approach, developed by the Center for Global Development (CGD), to incentivize local level actors to produce better learning outcomes. The COD focuses on paying for outcomes (such as levels of literacy and numeracy) that are independently verified, rather than paying for or micromanaging inputs (such as procuring textbooks or training teachers). The sequencing of disbursements is also important; payments are made *after delivery for outcomes* rather than *upfront for inputs*. It provides for payment commensurate with successful delivery – the better one performs, the more one gets. As such, COD potentially enables a higher level of public transparency and a greater opportunity for citizens to hold governments accountable for quality public service provision.

Our local COD intervention seeks to:

- test the extent to which paying for performance improves basic literacy and numeracy;
- vary levels of who receives the payment to test which option is more effective; and
- understand the effects of the intervention on public debate and citizen engagement.

The core component of our local COD design is to pay about USD 20 per pupil in **Grade 2 (or 3?, or both? TBD)** who performs well on the Uwezo Grade 2 level literacy and numeracy assessment. The 'offer' will be publicized at the beginning of the school year (around January/February) and children will be independently assessed towards the end of the school year (around November).

While current practice tends to target resources to national and district authorities, much of the action needed to improve basic learning needs to take place at the community level, and particularly in the interaction between teachers and pupils. Several studies suggest that incentivizing schools and particularly teachers to perform better may be more effective than increasing national budgets (Glewwe, Holla and Kremer, 2008; Muralidharan and Sundararaman 2011). We have therefore modified the original CGD design, which emphasizes a different way of providing aid to *national governments*, to instead create incentives for a varying level of *local level actors*. Indeed cash incentives are also likely to be more effective when targeted at those most at stake and where the payment amounts are likely to be considered significant. Additional advantages notwithstanding, the incentives for national ministries to reform practices due to a potential increase in COD Aid are comparatively weak.

Our research design will vary the local actors who will receive the payments, so as to test which level of targeting is most effective. Four variations will be used: a) districts, b) schools, c) teachers, d) communities, meaning teachers and parents. By teachers we mean the (Grade 2/3) teacher of the children who will be assessed, who in many cases is the same person charged to teach Grade 1.

Local Cash on Delivery intervention

Paid to	Arm	Experiment Arm
About USD 20 per child who does well on the Uwezo Grade 2 level assessment	5	<i>District focused:</i> USD 20 per child, paid to the district
	6	<i>School focused:</i> USD 20 per child, paid to the child's school
	7	<i>Teacher focused</i> USD 20 per child, paid to the child's teacher
	8	<i>Community focused:</i> USD 20 per child, USD 10 paid to the teacher and USD 10 to the child's parents or guardians
	C	<i>Control:</i> current practice, no action will be taken to modify government practice

Paying per child who does well introduces a potentially serious moral hazard, of schools or teachers focusing on a subset of pupils who they think can do well and ignoring poorly performing pupils. One way of handling this would be to only pay for the *additional* number of pupils who do well. But this option would require a baseline to be undertaken (which is both organizationally and financially costly), create complications regarding managing mid-year turnover, and could potentially reduce the incentive so much so that it ceases to be an effective driver. Instead, we propose that a minimum threshold pass level of the whole class need to be attained for the full COD payment to be made. The threshold would vary across districts, using the Uwezo district results as a reference, so as to take into account large geographical disparities. For example, if the most recent average Uwezo pass rate for a district was x%, x + 10% of the children would need to have passed for the payment to be made.

The COD approach is novel in Tanzania, and to our knowledge nothing of the sort has been tried before in education. Care will be taken to explain the experiment and in particular the basis of the assessment and payment carefully and in straightforward language, and to verify the understanding, so as to minimize misunderstanding. Popular communication materials will be developed and local communication channels used as appropriate, which being carefully not to undermine research design.

Twaweza, unlike the original CGD design, will commission an independent body to measure student learning outcomes, instead of auditing government-reported performance. We believe this provides more direct and reliable form of verification. Uwezo based assessment tools (though the actual materials will be different from those used in the actual annual Uwezo assessments). Uwezo has a demonstrated track record in undertaking large scale assessments at reasonable cost.

Intervention 3 – Combination of CGF and COD:

The COD approach is designed to be 'additional' or 'on top of' existing programs and budgets, and in effect create an incentive to make better use of those resources. But if existing resources are significantly inadequate or not disbursed, particularly at the school level, schools and teachers may simply be unable to take the actions necessary to achieve results for which they will be rewarded later. In other words, that a COD intervention can only reasonably expected to work after one has ensured basic inputs have been provided for. This concern was raised by several government leaders in a workshop organized by CGD and Embassy of Sweden in Dar es Salaam in April 2012.

This third combination intervention allows us to test this concern by combining CGF and COD. For practical purposes, not all variations of both experiments can be tested. The combined intervention will have two arms as follows:

Combined CGF and local COD intervention

Paid to	Arm	Experiment Arm
USD 10 CG to schools upfront and USD 20 per pupil in Grade 2/3 upon passing Uwezo assessment	9	<i>School focused:</i> USD 10 per pupil for all pupils to schools upfront, including information; plus USD 20 per Grade 2/3 pupil passing
	10	<i>Community focused:</i> USD 10 per pupil for all pupils to schools upfront, including information; plus USD 10 for child's teacher and USD 10 for child's parent/guardian per Grade 2/3 pupil upon passing
	C	<i>Control:</i> current practice, no action will be taken to modify government practice

General Research Aspects

All schools within chosen districts will be included in the RCT. In addition, the number of schools to be included will depend on the interest and support of government and donor partners. At the minimum, Twaweza aims to include twelve districts, which is about 1,000 to 1,200 schools. A major issue to be sorted involves sampling in relation to the district focused interventions. Indeed, while the policy questions and overall thrust of the experiments have been largely decided by Twaweza in consultation with partners, the final choice of variations and combinations deployed will be determined in consultation with independent researchers to ensure validity of research design.

Some of the other issues that require attention are:

- To reduce transport and administration costs we propose the experiments be geographically clustered by district. In each sampled district, we will only have one of the major treatments (CG, COD, or combination), while the schools within each district will be randomly assigned to each of the arms under this treatment. How to compensate schools that are not assigned any 'treatment' without affecting research design will also need to be considered.
- Qualitative studies: Twaweza will request the research team to conduct qualitative studies in parallel to the RCT to ensure results can be explained in the local context and to help unpack the 'black box' of why something happened or not. Potential issues include use of different types of information, effect of an influx of financial support for only certain teachers, barriers from existing local power structures, transfer of teachers to where there are incentives.
- Concerns about how COD money, especially to the school level, is spent: while the disbursements are meant to be 'hands off' regarding how it should be spent, there are worries since local governance and accountability arrangements are not always robust. A potential response is insisting on full disclosure and public transparency.
- Two kinds of substitution of funds: 1) the amount of money parents usually contribute to the school or use for their children's education may decrease, thus reducing the total amount of money used for school. 2) As the capitation grant intervention from Twaweza is in addition to the CG from the government, local districts may attempt to 'even out' the money each school receives in total by disbursing less to treatment schools. This would reduce the control effect of the RCT.

3. Policy Considerations

These experiments are designed to produce rigorous evidence on how to turn investments in education into improved learning, a challenge facing educationalists and policy makers in both Tanzania and globally. This work is expected to engage with and inform policy and practice, and therefore the possibilities and politics of turning evidence into action need to be taken into serious consideration from the outset of the research design.

Overall, undertaking the capitation grant flow and local COD interventions in parallel will enable policy makers to compare the results of the existing but fully implemented policy of capitation grant interventions with that of the local COD innovation. Put differently, this approach allows us to compare the impact on learning of an input based intervention (capitation grants) with that of an outcome based incentive intervention (local COD). The third intervention tests the impact of doing both at the same time. Lessons generated from these pilots and their rigorous evaluation can provide a solid evidence base for scaling up interventions across the country, and contribute to the global knowledge base for improving learning.

Pilot projects carry the inherent risk of failing to be scaled up or sustained over time. Many aspects can affect the failure of pilots to improve policy and practice; we have identified five to be among the most important, and seek to have this design take each into account as follows:

- *Interventions are parachuted in from up above, without participation of relevant actors:* While safeguarding the integrity and independence of the intervention, we will seek to involve all key actors throughout the process, from national and local government, the teachers' union, civil society, think tanks and researchers, and donors. See next section for list of partners.
- *Interventions remain technocratic and findings sit on shelves:* In addition to involving the partners noted above, we will focus in on district and community level actors, including in particular MPs, who are well placed to motivate local action as well as create 'middle-down' and 'middle-up' pressures for change. For this reason we may consider doing some purposive sampling involving key MPs, whose areas may not then technically part of the RCT analysis. We will publicize findings and lessons widely in accessible language, including through the media, so as to stimulate public debate.
- *Intervention design is too elaborate; requires too high level skills and unrealistic commitment:* Both the capitation grant and local COD designs are simple in design, and can be implemented nationwide with limited additional preparation of capacity; the independent Uwezo assessment has already demonstrated it can be scaled up nationwide.
- *Unit costs are too high to be scaled up and sustained over time:* Nationwide full coverage of the capitation grant intervention (about USD 90-95 million per year) is in fact existing policy and could be paid from reprioritizing existing budgets. The local COD if scaled up for each Grade 2 or 3 child nationwide would cost about USD 30 million per year, *if every child were to pass*, or more realistically USD 15 million estimating about 50% would pass initially. The elegance of the COD also is that one only pays for results. No results=no payment.
- *Interventions are not properly evaluated and lack sufficient independence and rigor:* From the outset we have engaged globally renowned, independent experts who will use RCT methodology complemented by other qualitative methods, so as to ensure generation of high quality evidence.

4. Research and Implementation partners

Twaweza will provide the overall leadership and coordination of these interventions, working in close cooperation with the partners noted below. A staff member will be recruited to manage the different aspects and partners of the exercise, and various Twaweza units will provide support on both implementation and learning aspects in accordance with their comparative advantage. The key partners to be involved are:

Government of Tanzania: supports the idea in principle. The government is exploring the best way to disburse full capitation grant to schools and is open to exploring the Cash on Delivery innovation. President Kikwete has personally engaged with CGD and Twaweza on COD and encouraged it be implemented in Tanzania.

State House: The President's personal commitment and political leadership will be called upon to guide and promote the interventions.

Prime Minister's Office (Regional Administration and Local Government): PMO-RALG oversees management of schools, including flow of information and use and monitoring of funds.

Ministry of Education and Vocational Training: MOEVT and its key institutions will be engaged on policy implications.

Ministry of Finance and Economic Affairs: MOFEA is responsible for overall allocation of funds and to ensure alignment with policy goals and value for money

COSTECH: COSTECH will review research design to ensure scientific rigor, issue research permit and help promote evidence based policy-making.

Tanzania Teachers' Union (TTU): TTU will provide important inputs from the teachers' perspective, in research design, participation in the study and dissemination of results.

Centre for Global Development (CGD): CGD, the original developer of the COD idea, will support concept development and learning, and support policy engagement globally and in Tanzania.

Joint Poverty Action Lab/Innovations for Poverty Action: JPAL, based at MIT, will lead the research design and impact evaluation under the guidance of Abhijit Banerjee, with Isaac Mbiti as lead researcher. We will also seek to engage with Karthik Muralidharan (UCSD), Ted Miguel (CEGA), Annie Duflo (IPA) and others to advise on key steps of the project. JPAL will develop a fuller research design and budget, and help disseminate results.

Uwezo: Uwezo, a semi-autonomous initiative of Twaweza, will provide the tools to assess literacy and numeracy that will be used to measure impact, under guidance of independent researchers.

Donors: Various Donors, including DFID, SIDA and the World Bank, are interested in strengthening public service provision transparency and accountability, especially in improving learning outcomes. In addition to policy engagement and providing core support to Twaweza, donors may provide additional support to Twaweza, the government and/or research partners as needed.

Annex A: Summary of Literature Review (work in progress, subject to review)

This continuously updated review of current literature is conducted through careful reading of related books and published articles. In this section, we review recent quantitative – RCTs and retrospective correlation – and qualitative studies that seek to evaluate whether interventions improve student learning outcomes, usually measured by test performance. It should be noted that many earlier studies only looked at indirect indicators, such as teacher attendance or student enrollment, not actual learning outcomes. Unless otherwise noted, we focus on studies that measured learning outcomes, not inputs or indirect measures of success. Approximately 80% of the 34 articles included are published in a peer reviewed journal. The book reviewed is *Poor Economics* by Abhijit Banerjee and Esther Duflo (2011). The main categories of intervention evaluation study results, with the most effective listed first, can be divided into financial incentives, pedagogical change, local empowerment, and physical inputs. Within each category, randomized control trial studies are discussed first, followed by other quantitative and qualitative approaches.

Financial Incentives

Many studies look at the effects of different types of financial incentives that try to improve student learning. Kremer and Holla (2008), in a literature review study, suggest that strategies to reduce out-of-pocket education costs consistently increase school participation and/or test scores. Cash incentives targeting outcomes are found to help schools, public or private, become more efficient at providing education to a larger number of students in a retrospective study carried out in two urban Kenyan slums (Oketch, Mutisya, et al. 2010). Turrent and Oketch (2009) in a literature review suggest we need to find mechanisms to fund education – one of which may be COD – that work in complex situations, such as post-conflict areas, which have been traditionally refused donor funding due to accountability issues. It is important to note that incentive intervention details differ substantially; as such, results cannot be easily extrapolated. For example, some programs reward at a school level, others at a teacher level; some pay larger incentives; some calculate the test scores differently. In general, however, the research in types of education financial incentives can be separated into merit scholarships, performance-linked-pay (PLP), and conditional cash transfers (CCT).

Merit Scholarships: Merit scholarships are offered to the best performing students, sometimes within a specific category, such as sex or socio-economic condition. Such scholarships seem to be one of the more cost-effective ways to improve test scores and teacher attendance (Kremer and Holla 2009). An often cited publication is on a randomized evaluation study on merit scholarships for primary school girls in two Kenyan districts (Kremer, Miguel and Thornton 2009). The scholarships were awarded to grade six girls who scored in the top 15% on district wide exams. The winning girls received two academic years' worth of school fees and supplies, in addition to a well-publicized awards ceremony. The study found that even after incentives were removed, the test score gains remained large in the following year, which may reflect real learning (Kremer, Miguel and Thornton 2009). They found no adverse changes in student attitudes towards school or an increase in for-test-only study preparation. Although the program only rewarded girls, the spillover effects also benefited boys and weaker female students who had no chance of obtaining a scholarship (Kremer, Miguel and Thornton 2009).

Conditional Cash Transfers (CCT): CCT programs, generally social welfare initiatives implemented by the government, transfer payments to families to support their livelihoods. The programs come in different forms: for example, school education vouchers, monthly family support payments, or health-related transfers. While the initial recipients are awarded randomly, continuation of these payments are usually conditional on the family performing certain responsibilities, such as parents taking their children for vaccinations or the children passing school tests. Due to the lottery nature of

CCT programs, many evaluations used the RCT method. Results usually indicate that CCT programs have slight to significant impacts on the test scores and learning of the children from participating families (Kremer and Holla 2009; Angrist, Bettinger and Kremer 2006). The most well-known and large scale CCTs are the Progresa (later renamed Oportunidades) program in Mexico and the Bolsa Familia program in Brazil. We also look at the Paces lottery private school voucher program from Colombia.

The Progresa program, evaluated in one RCT study for rural areas, increased child enrolment in schools and was especially effective for encouraging older children to stay in school (Attanasio, Meghir and Santiago 2005). The largest increase in enrolment was among girls who had completed grade 6 (Schultz 2003). The research results were however, focused on enrolment, not learning outcomes. It was also found that deferring payment of CCTs to coincide with the time for school fee payment led to a larger impact on subsequent enrolment than evenly spaced transfers throughout the year, may be due to time-inconsistent preferences or savings constraints (Kremer and Holla 2009), which may be relevant to Twaweza’s intervention design.

The Colombian education vouchers program Paces, one of the largest of its kind, is a CCT program conditional on grade maintenance by students. Initially distributed by lottery, it aimed to encourage children from poor neighborhoods to attend private schools. Ultimately, the program led to a substantial increase in both attendance and test scores in participants (Angrist, Bettinger and Kremer 2006). The voucher winners, compared to those who were not chosen in the lottery, also had much higher high-school graduation rates, attended more expensive private schools, and appear to have learned more (Angrist, Bettinger and Kremer 2006). As continuation of receiving vouchers was dependent on a passing grade, there was a large incentive for lottery winners to devote more efforts at school (Angrist, Bettinger and Bloom, et al. 2002).

Performance Linked Pay (PLP): this category is most similar to the Twaweza COD initiative. Interventions in PLP vary widely in design and impact, however, Kremer and Holla (2009) conclude that there is a general trend of positive, though not always persistent post-program, impacts.

In 3 out of 4 randomized control trial studies reviewed, students in the treatment schools performed slightly to significantly better than those in the control school (Muralidharan and Sundararaman 2011; Glewwe, Ilias and Kremer 2010; Duflo and Hanna 2005; Olken, Onishi and Wong 2012). In the study where students did not perform any better, the incentives did not specifically target test score outcomes. One study found that performance based bonus payments to teachers are more cost-effective at increasing student test scores, compared to spending similar amounts on schooling inputs (Muralidharan and Sundararaman 2011). A summary of the studies is shown in the table below; details are described in the following paragraphs.

	Intervention	Comparison to Twaweza
Muralidharan and Sundararaman 2011	Two types of teacher performance pay: group bonuses based on school performance and individual bonuses based on teacher performance. Average bonus calibrated to be around 3 percent of a typical teacher’s annual salary.	<ul style="list-style-type: none"> Twaweza’s COD is calculated based on each child that passes the Uwezo assessment, not group performance.
Glewwe, Ilias and Kremer 2010	In-kind prizes awarded to grades 4 to 8 teachers based on average student test performance in annual district exams. Collective incentive. Test score improvements during program, not after. No real improvement in pedagogy; suspicion of “teaching to the test.”	<ul style="list-style-type: none"> Twaweza’s COD awards cash and targets different groups concurrently.

Duflo and Hanna 2005	Teacher's salary based on attendance: a base salary and bonus/penalty per extra day/absence (Rs500-1300/month). Control schools teachers paid flat rate (Rs1000). 30% more instruction time; teacher absence halved (44% to 22%). treatment schools' test scores significantly higher than comparison schools'.	<ul style="list-style-type: none"> • Twaweza's COD is based on test scores, not attendance. It is a lump sum that doesn't vary.
Olken, Onishi and Wong 2012	Conditional on performance for subsequent grants, villages received block grants to improve health and education indicators. Control villages either received identical, but unconditional, block grants or no grants. Impacts found on health, but not education (enrollment and attendance, not test scores).	<ul style="list-style-type: none"> • Twaweza's COD is not a block grant and does not target health outcomes.

A RCT teacher incentive program in Kenya showed that students in program schools attain higher scores on certain exams, but only for the duration of the program (Glewwe, Ilias and Kremer 2010). The researchers suspect gains are mainly due to pre-test training sessions, and not improved teacher attendance, homework assignment or pedagogy (Glewwe, Ilias and Kremer 2010). Subjects that were not linked to teacher incentives did not improve, while students improved significantly on multiple-choice questions compared to fill-in-the-blank ones (Glewwe, Ilias and Kremer 2010). No effects on drop-out rates were found but a higher share of students in program schools took the test, which shows that teachers were not trying to keep the weaker students from taking the exam (Glewwe, Ilias and Kremer 2010). This intervention shows the importance of choosing the incentive indicator correctly to ensure we incentivize the desired outcome. The authors also suggest that the incentive amount may have been too low and that the teachers knew the program would not operate beyond two years (Glewwe, Ilias and Kremer 2010).

The RCT and qualitative evaluation of a PLP intervention in Andhra Pradesh, India indicate that not only did student test scores improve, the teachers were also very supportive of the program (Muralidharan and Sundararaman 2011a,b). Students in the treatment schools performed significantly better than those in the control schools at the end of two years (Muralidharan and Sundararaman 2011b). School-level group incentives and teacher-level individual incentives performed equally well in the first year, but the individual incentive schools were more significant at the end of two years (Muralidharan and Sundararaman 2011b). The main causes were not increased teacher attendance, but more effective teaching efforts (Muralidharan and Sundararaman 2011b). In a parallel study, they provided an extra contract teacher or a cash grant for school materials to two groups of randomly chosen schools. The PLP program schools outperformed both of the input-related cash grant and contract teacher treatment arms. They found that performance based bonus payments to teachers are significantly most cost-effective at increasing student test scores, compared to spending similar amounts on schooling inputs (Muralidharan and Sundararaman 2011b).

The qualitative study of the same intervention found that teachers were being de-motivated previous to the program because there were no better career prospects dependent on performance (Muralidharan and Sundararaman 2011a). They also trusted the integrity of the program, run by an NGO, and thought the assessment tools were an appropriate and fair measure of student learning (Muralidharan and Sundararaman 2011a). The potential worries – such as teachers feeling that evaluating only on test scores would limit their functioning as teachers; holding teachers as the only one accountable to student learning when other factors like household characteristics are just as important; or competitiveness between teachers would reduce workplace collegiality; mistrust of administrations, and so on – did not appear in this qualitative study result (Muralidharan and Sundararaman 2011a). The researchers suggested that PLP programs should frame the program less in terms of “school accountability” and more in terms of “teacher recognition” so that the rewards

seem consistent with the teachers' own notions of good professional conduct (Muralidharan and Sundararaman 2011a).

Another intervention in India linked pay bonuses to teacher attendance, not test scores, tracked by using digitally dated self-taken photos by the teachers with students (Duflo and Hanna 2005). The RCT evaluation by Duflo and Hanna (2005) indicates that students benefited from 30% more instruction time and significantly higher test scores than comparison schools (Duflo and Hanna 2005). Students who originally had higher test scores gained more compared to less competent students. Teachers reported feeling empowered by the ability to earn more from better performance and that the community burdened them less with other responsibilities once it was known that absences were penalized (Duflo and Hanna 2005).

Olken, Onishi and Wong (2012) find no effects for a large-scale Indonesian experimental incentive program on schooling outcomes, although only enrollment and attendance, not test scores, were considered. The incentive, however, took the form of a village block grant while the performance was measured through an index consisting of 12 measures, some of which were health outcomes and some education outcomes (Olken, Onishi and Wong 2012).

Pedagogical Change

Pedagogical changes refer to changes in ways of teaching, structuring classes, or communicating information. Specifically, strategies that target weaker students, hire contract teachers, change the curriculum, and communicate report cards, seem to have, to some extent, positive effects on student learning outcomes.

Overall, targeting weaker students, and, related, the hiring of contract teachers, seem to have the largest effect. Setting up class structures to ensure that the weakest students in the class understand the basics of what is being taught is a strategy that seems to consistently and significantly increase the over class evaluation results. Different methods include tracking students (separating them into classes/groups according to their abilities), remedial classes after school with teachers or volunteers, and technology-assisted learning. This category of literature tends to suggest that weaker students cannot take advantage of additional physical inputs because of their already low ability to read, write, and learn (Glewwe, Kremer and Moulin 2009; Kremer and Holla 2009).

Tracking is a strategy that does not need any extra input, provided there are already enough teachers to divide classes, where RCT impact evaluations show that students of all levels perform better than in non-tracked schools (Duflo, Dupas and Kremer 2012). Separating students allow teachers to tailor their curriculum to those who are weaker, enabling them to learn at their own pace; hence being advantages to both high and low achieving students (Kremer and Holla 2009). There are concerns that teachers will tend to seek high achieving students if there are incentives for pay bonuses due to test scores, or simply disregard those students who are too weak to pass (Duflo, Dupas and Kremer 2012). Suggestions are that teachers' performances can be incentivized on a value-added basis, rather than based on end line scores (Duflo, Dupas and Kremer 2012).

Remedial classes, outside of school or during class time, taught by teachers or volunteers also can improve learning outcomes. One RCT intervention in India recruited village volunteers to run reading camps, which turned out to be "remarkably effective" in improving literacy (Banerjee et al 2010). Weaker students can also be taken out of class time to work in smaller groups with other teachers. Although another pilot with Banerjee et al (2007) found that the large gains in test score improvements were short-lived, only about a year after program end, some of the effects persisted over time.

One retrospective evaluation in rural Colombia found that providing nine different models of educational materials for municipalities to choose which suited their context most significantly lowered student dropout rates and test failure rates (Rodriguez, Sanchez and Armenta 2010). This suggests that the nationally approved curriculum sometimes is biased towards an urban, elite population (Glewwe, Kremer and Moulin 2009).

A few interventions have experimented with the possibility of hiring contract teachers – teachers who are usually less experienced, local to the school area, willing to work for lower pay, and not under national teacher unions (Kremer and Holla 2009). A Kenyan RCT intervention provided funds for the Parent-Teacher Associations to hire an extra contract teacher. The results seem to indicate that this strategy increases student learning, reduces the school financial burden of teacher salaries, and motivates previously hired permanent teachers (Duflo, Dupas and Kremer 2012). The researchers suggest, however, that unless similar decentralization initiatives are combined with institutional reforms to allow local level democratic control, especially by parents or the community, local state agents may capture rents from such programs (Duflo, Dupas and Kremer 2012).

Lastly, report cards that show student and school performance compared to other students and schools seem to have no direct effect on student learning, but some side-effects. One study by Muralidharan and Sundararaman in India indicate that the report cards, although they raised awareness, didn't have any impact on learning until coupled with performance-linked pay (2010). At the same time, Andrabi, Das and Khwaja show that Pakistani parents tend to overestimate the quality of their child's school (2009). The report card which compared different schools gave parents a tool to contrast education quality through third-party verification (Andrabi, Das and Khwaja 2009). The low performing private schools increased their investment in learning material and less break time between classes with limited fee changes; the government schools also increased learning outcomes by a smaller, but significant, amount (Andrabi, Das and Khwaja 2009). One exception may be an intervention in Madagascar where the information provided along with the report card was coupled with an introduction of parental monitoring through school meetings. While the interventions at the district and sub district level had minimal effects, the interventions at the school level significantly improved teacher behavior, test scores, and attendance (Lassibille et al 2010).

Local Empowerment

Interventions that can be broadly categorized as local empowerment come in different forms such as parent-teacher associations, village education committees, parent training programs, school-related management reforms, and media campaigns. The results from these studies often contradict as intervention designs vary widely dependent on reasons such as cultural context, combination of incentives, and reputability of the implementation partner.

Attempts at eliciting pressure from parents to hold teachers and schools accountable for improving student learning outcomes have, in general, not resulted in much impact. Variants of these RCT interventions usually test whether information alone about the state of their children's learning is enough to prompt action from parents (Kremer and Holla 2009). Results from randomized evaluations are usually negative, unless combined with financial incentives, increases in management power, such as hiring and firing teachers, or showing alternative ways to work around the formal education system (Banerjee, Banerji, et al. 2010).

Some studies add in a parent empowerment training component to see if parents would take the information they have learned to monitor schools and demand better services. RCT interventions which only provided information, but not action-oriented suggestions, have generally failed to produce statistically significant improvements (Duflo, Dupas and Kremer 2011; Blimpo and Evans 2011). A common reason cited for such outcomes in literature review is that parents lack agency in

thinking that they can hold schools accountable (Banerjee and Duflo 2006) or they simply do not pay much attention to the role of public participation or collective action (Banerjee, Banerji, et al. 2007). One survey study in India found that 92% of parents were unaware of the existence of a Village Education Committee which is theoretically the venue to press for better education service (Banerjee, Banerji, et al. 2007). Another suggestion is that often less educated parents cannot verify for themselves how much their children are learning. A RCT intervention studying the effects of grants and school-management training in Gambia found that while student test results were poor, over 90% of the parents were satisfied with the schools' and their children's performance (Blimpo and Evans 2011).

Positive results, on the other hand, include involving parents in using student assessment tools and media campaigns on school funding. In two treatment arms of a RCT intervention in India, parents and other villagers were taught how to use the learning level assessment tools on school children (Banerjee, Banerji, et al. 2007). The poor results seem to have prompted more reactions from the parents, greatly increased children learning when, under suggestion from the implementing NGO, villagers volunteered their time to teach the weaker children basic literacy in remedial reading camps (Banerjee, Banerji, et al. 2007). The authors believe that the offering of an alternative outside lobbying the formal bureaucratic system empowered the villagers.

In addition, a media campaign retrospective study in Uganda is often cited as a success in demonstrating the power of citizen agency and public pressure. The authors find that providing information to the media on the amount of capitation grant to be received by schools dramatically increased the amount actually received (Reinikka and Svensson 2005). Schools that are more extensively exposed to the newspaper campaign experience a significantly larger reduction in local capture of funds; some schools received 82% of its entitlements (Reinikka and Svensson 2005). Critics of the study note that Reinikka and Svensson did not study whether these funds improve schooling outcomes (Hubbard 2007). Yet the role of information in ensuring full delivery of funding is highly relevant to the Twaweza capitation grant intervention.

Physical inputs

Amongst the randomized trial literature is an almost uniform consensus that physical inputs, such as textbooks or flip charts, have little to no direct impact on learning outcomes (Glewwe, Kremer and Moulin 2009; Glewwe, Kremer, et al. 2004; Banerjee, Cole, et al. 2007). Non-experimental multivariate analysis from Uwezo's 2010 data shows a similarly weak relationship between inputs and outcomes. In rural Kenya, the additional textbooks provided in one RCT intervention only seemed to have an effect on those students who were already more competent (Glewwe, Kremer and Moulin 2009). The researchers suggested that due to the non-native language textbooks, weaker students were not able to take advantage of the extra learning material (Glewwe, Kremer and Moulin 2009). Another literature review study suggested that "elite-oriented curricula" that privileges educational levels of urban students rather than rural students contributes to this outcome (Kremer and Holla 2009). Extra textbooks are also ineffective when faced with consistent teacher absence to the tune of 1 in 5 during unannounced visits (Kremer and Holla 2009). A randomized intervention in India suggests that "for many children, neither more inputs nor an extra day makes much of a difference because what is being taught in class is too hard for them (Banerjee, Cole, et al. 2007)." One possible exception is school meals and deworming programs which promote nutritional intake of school age children. While the direct retrospective testing results do not show large increases in literacy and numeracy, there are small, but statistically significant results, on learning and teaching improvements (Aturupane, Glewwe and Wisniewski 2011; Banerjee and Duflo 2006). A literature review on RCT interventions suggests that nutrition-related programs are also one of the most cost-effective programs to implement (Kremer 2003).

The proposed design of the Twaweza COD intervention is close to the teacher incentive intervention described in Muralidharan and Sundararaman (2011): a large-scale randomized evaluation of a performance linked pay program for teachers implemented in the Indian state of Andhra Pradesh. There were two types of teacher performance linked pay – group bonuses based on school performance and individual bonuses based on teacher performance – where the average bonus was calibrated to be around 3 percent of a typical teacher’s annual salary. At the end of 2 year intervention, students in incentive schools performed significantly better than those in control schools. Both school-level group incentives and teacher-level individual incentives performed equally well in the first year, but the individual incentive schools outperformed the group incentive schools at the end.

The Capitation Grant component of Twaweza focuses on directly channeling funding to schools, which could be used for physical inputs by the schools’ discretion. While physical inputs may not directly increase test scores, other non-randomized evaluations also suggest that basic levels of learning materials (e.g. chalk, books, and pens) are needed (Kremer and Holla 2009). In addition, there seem to be no existing studies that evaluate the impact of channeling capitation grants directly to schools. Such evaluation is worthwhile as this is closest to current policy. The evidence for policy change to an output-based strategy (COD) will be greatly strengthened if the output-based arms outperform a well-implemented capitation grant treatment.

Works Cited

- Andrabi, Tahir, Jishnu Das, and Asim Ijaz Khwaja. "Report Cards: The Impact of Providing School and Child Test-scores on Educational Markets." *Bureau for Research and Economic Analysis of Development, Working Paper No. 226*, 2009.
- Angrist, Joshua, Eric Bettinger, and Michael Kremer. "Long-Term Educational Consequences of Secondary School Vouchers: Evidence from Administrative Records in Colombia." *The American Economic Review*, 2006: 847-862.
- Angrist, Joshua, Eric Bettinger, Erik Bloom, Elizabeth King, and Michael Kremer. "Vouchers for Private Schooling in Colombia: Evidence from a Randomized Natural Experiment." *The American Economic Review*, 2002: 1535-1558.
- Attanasio, Orazio, Costas Meghir, and Ana Santiago. "Education Choices In Mexico: Using A Structural Model And A Randomized Experiment To Evaluate Progresa." *Centre for the Evaluation of Development Policies, Report*, 2005.
- Aturupane, Harsha, Paul Glewwe, and Suzanne Wisniewski. "The impact of school quality, socioeconomic factors, and child health on students’ academic performance: evidence from Sri Lankan primary schools." *Education Economics*, 2011: 1–37.
- Banerjee, Abhijit, and Esther Duflo. "Addressing Absence." *Journal of Economic Perspectives* 20, no. 1 (2006): 117–132.
- . *Poor Economics: A Radical Rethinking of the Way to Fight Global Poverty*. New York: Perseus Books Group, 2011.
- Banerjee, Abhijit, et al. "Can Information Campaigns Raise Awareness and Local Participation in Primary Education?" *Economic and Political Weekly* 42, no. 15 (2007): 1365-1372.
- Banerjee, Abhijit, Rukmini Banerji, Esther Duflo, Rachel Glennerster, and Stuti Khemani. "Pitfalls of Participatory Programs: Evidence from a Randomized Evaluation in Education in India." *American Economic Journal: Economic Policy* 2, no. 1 (2010): 1–30.

- Banerjee, Abhijit, Shawn Cole, Esther Duflo, and Leigh Linden. "Remedying Education: Evidence from Two Randomized Experiments in India." *The Quarterly Journal of Economics*, 2007: 1235-1264.
- Blimpo, Moussa P, and David K. Evans. *Experiment, School-Based Management and Educational Outcomes: Lessons from a Randomized Field*. November 2011.
- Duflo, Esther, and Rema Hanna. "Monitoring Works: Getting Teachers to Come to School." *National Bureau of Economic Research, Working Paper No. 11880*, 2005.
- Duflo, Esther, Pascaline Dupas, and Michael Kremer. "Peer Effects, Teacher Incentives, and the Impact of Tracking: Evidence from a Randomized Evaluation in Kenya." *American Economic Review* 101, 2011: 1739–1774.
- Duflo, Esther, Pascaline Dupas, and Michael Kremer. "School Governance, Teacher Incentives, and Pupil-Teacher Ratios: Experimental Evidence from Kenyan Primary Schools." *National Bureau of Economics Research Working Paper No. 17939*. Cambridge, MA, March 12, 2012.
- Glewwe, Paul. "Schools and Skills in Developing Countries: Education Policies and Socioeconomic Outcomes." *Journal of Economic Literature* 40 (2002): 436–482.
- Glewwe, Paul, Alaka Holla, and Michael Kremer. "Teacher Incentives in the Developing World." September 11, 2008.
http://www.economics.harvard.edu/faculty/kremer/files/KntchinV9_080915.pdf (accessed April 23, 2012).
- Glewwe, Paul, and Eugenie Maïga. "The Impacts of School Management Reforms in Madagascar: Do the Impacts Vary by Teacher Type?" *Journal of Development Effectiveness* 3, no. 4 (2011): 435-469.
- Glewwe, Paul, Michael Kremer, and Sylvie Moulin. "Many Children Left Behind? Textbooks and Test Scores in Kenya." *American Economic Journal: Applied Economics* 1, no. 1 (2009): 112–35.
- Glewwe, Paul, Michael Kremer, Sylvie Moulin, and Eric Zitzewitz. "Retrospective vs. prospective analyses of school inputs: the case of flip charts in Kenya." *Journal of Development Economics*, no. 74 (2004): 251– 268.
- Glewwe, Paul, Nauman Ilias, and Michael Kremer. "Teacher Incentives." *American Economic Journal: Applied Economics*, 2010: 205–227.
- Hubbard, Paul. "Putting the Power of Transparency in Context: Information's Role in Reducing Corruption in Uganda's Education Sector." *Center for Global Development, Working Paper Number 136*, 2007.
- Kremer, Michael. "Randomized Evaluations of Educational Programs in Developing Countries: Some Lessons." *The American Economic Review* 93, no. 2 (2003): 102-106.
- Kremer, Michael, and Alaka Holla. "Improving Education in the Developing World: What Have We Learned from Randomized Evaluations?" *Annual Review of Economics* 1 (2009): 513-542.
- Kremer, Michael, Edward Miguel, and Rebecca Thornton. "Incentives To Learn." *The Review of Economics and Statistics* 91, no. 3 (2009): 437–456.
- Lassibille, Gerard, Jee-Peng Tan, Cornelia Jesse, and Trang Van Nguyen. "Managing for Results in Primary Education in Madagascar: Evaluating the Impact of Selected Workflow Interventions." *World Bank Economic Review* 24, no. 2 (2010): 303-329.
- Muralidharan, Karthik, and Venkatesh Sundararaman. "Teacher opinions on performance pay: Evidence from India." *Economics of Education Review*, no. 30 (2011): 394–403.

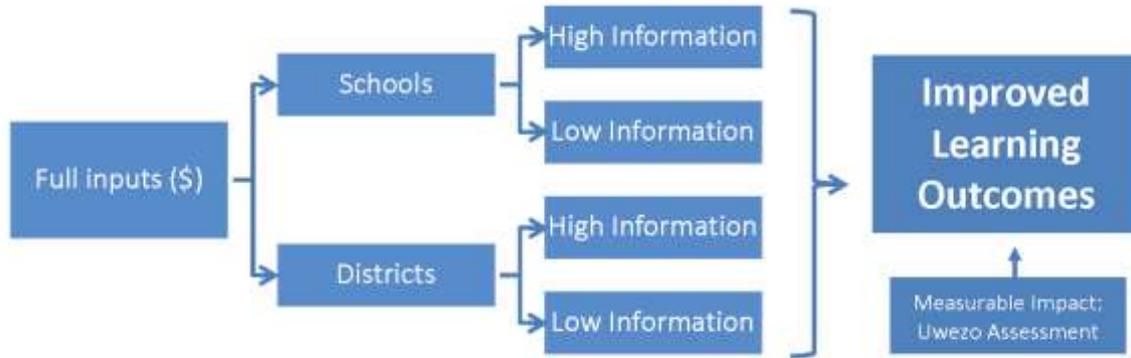
- Muralidharan, Karthik, and Venkatesh Sundararaman. "The Impact Of Diagnostic Feedback To Teachers On Student Learning: Experimental Evidence From India." *The Economic Journal*, no. 120 (2010): 187–203.
- Muralidharan, Karthik, and Venkatesh) Sundararaman. "Teacher Performance Pay: Experimental Evidence from India." *Journal of Political Economy* 119, no. 1 (2011): 39-77.
- Oketch, Moses, and Moses Ngware. "Free primary education still excludes the poorest of the poor in urban Kenya." *Development in Practice* 603-610 (2010): 603-610.
- Oketch, Moses, Maurice Mutisya, Moses Ngware, Alex C. Ezeh, and Charles Epari. "Free primary education policy and pupil school mobility in urban Kenya." *International Journal of Educational Research* 49 (2010): 173–183.
- Olken, Benjamin A., Junko Onishi, and Susan Wong. "Should Aid Reward Performance? Evidence from a Field Experiment on Health and Education in Indonesia." *NBER Working Paper No. 17892*, 2012.
- Reinikka, Ritva, and Jakob Svensson. "Fighting Corruption to Improve Schooling: Evidence from a Newspaper Campaign in Uganda." *Journal of the European Economic Association* 3 (2005): 1–9.
- Rodriguez, Catherine, Fabio Sanchez, and Armando Armenta. "Do Interventions at School Level Improve Educational Outcomes? Evidence from a Rural Program in Colombia." *World Development* 38, no. 3 (2010): 415–428.
- Schultz, Paul. "School Subsidies for the Poor: Evaluating the Mexican Progresa Poverty Program." *Journal of Development Economics*, 2003.
- Turrent, Victoria, and Moses Oketch. "Financing universal primary education: An analysis of official development assistance in fragile states." *International Journal of Educational Development*, no. 29 (2009): 357–365.
- Zhao, Meng, and Paul Glewwe. "What determines basic school attainment in developing countries? Evidence from rural China." *Economics of Education Review* 29 (2010): 451–460.

Graphical representations

(Draft-To be tweaked)

The following graphics summarize the variations of the treatment arms the different COD and CG interventions in the Twaweza pilots:

Capitation Grant Flow Intervention



Local Cash on Delivery Intervention



Combination: CG and COD

