



Annual research digest 2017-18

Evidence from the developing world

CENTRE FOR EDUCATION ECONOMICS



Annual research digest 2017-18

The Centre for Education Economics wishes to thank Cambridge Assessment for their sponsorship, without which this project would not have been possible.



About the editor

Lee Crawford (@leecrawford) is a Strategic Advisor with the Ministry of Education in Rwanda and the Tony Blair Institute for Global Change, and a PhD candidate in economics at the University of Sussex. He has worked on global development and education policy as a researcher, consultant, and civil servant for the past ten years. He has published peer-reviewed articles in the *Journal of African Economies and Development Policy Review* writes a popular blog www.rovingbandit.com



First edition published 2018 by the Centre for Education Economics CIC

Copyright © Centre for Education Economics CIC 2018

ISBN 978-1-5136-4131-7

The moral rights of the author have been asserted.

All rights reserved. No part of this publication may be reproduced, stored in or introduced into a retrieval system, or transmitted in any form, or by any means (electronic, mechanical, photocopy, recording, or otherwise) without the prior permission of the publisher. Any person who does any unauthorised act in relation to this publication may be liable to criminal prosecution and civil claims for damages.

A CIP catalogue record for this book is available from the British Library.
The publisher accepts no responsibility for any error in detail, inaccuracy or judgement whatsoever.

Every care has been taken that all information was correct at the time of going to press.

Typeset by Konstant and New Designs, London.

Printed and bound by Lightning Source Ltd.

About the authors

Barbara Bruns is a Visiting Fellow at the Center for Global Development and former lead education economist at the World Bank, where she specialised in Latin American education and rigorous evaluation of education programmes. As the first manager of the \$14 million Spanish Impact Evaluation Fund (SIEF) at the World Bank from 2007 to 2009, she oversaw the launch of more than fifty rigorous impact evaluations of health, education, and social protection programmes. Barbara also headed the EFA Fast Track Initiative (now Global Partnership for Education) from 2002-2004 and served on the 2003 Education Task Force appointed by the UN Secretary General.

Celeste Carano is a Governance Advisor to the Ministry of Education in Liberia with the Tony Blair Institute for Global Change. She previously led fundraising and strategy for More Than Me, a Partnership Schools provider in Liberia, and worked on network growth at Teach For All.

Jishnu Das is a Lead Economist in the Development Research Group at The World Bank and a Senior Visiting Fellow at The Center for Policy Research, New Delhi. Jishnu's work focuses on health and education in low and middle-income countries and he was member of the core team for the World Development Report on Gender and Development, 2011. He was the Flander's Visiting Professor at McGill University in 2015, received

the George Bereday Award from the Comparative and International Education Society, the Stockholm Challenge Award for the best ICT project in the public administration category in 2006, and the Research Academy award from the World Bank in 2017 and 2013. He has worked in India, Pakistan, Nepal, Paraguay, Kenya and Zambia.

Aidan Eyakuze (@aeyakuze) is the Executive Director of Twaweza East Africa. Twaweza works to enable children to learn, citizens to exercise agency and governments to be more open and responsive in Tanzania, Kenya and Uganda. He is on the global Steering Committee of the Open Government Partnership (OGP) and the Board of the Global Partnership for Sustainable Development Data (GPSDD). Before joining Twaweza, Aidan was Associate Regional Director of the Society for International Development (SID) and Head of the SID Tanzania office. Aidan is an Archbishop Desmond Tutu Leadership Fellow and a Board Member of the African Leadership Institute (South Africa).

Susannah Hares (@Fihi_maFihi) is Senior Policy Fellow and Global Education Director at the Center for Global Development (CGD). Before joining CGD, she spent seven years as Ark's International Director, where she was responsible for strategy, operations and programmes in India, sub-Saharan Africa and Eastern Europe, and for Ark's international ventures including the Education Partnerships Group, Global Schools Forum and Peepul. Prior to this Susannah spent two years as Associate Director, Global Community Partnerships, at Barclays PLC. She started her career in Kenya, working for African health NGO AMREF. She has served on the boards and advisory groups of various international education organisations and ventures.

She holds a MA(Econ) in Political Development from the University of Manchester.

Lant Pritchett is Professor of the Practice of International Development at Harvard's Kennedy School of Government, and Senior Fellow at the Center for Global Development. He is currently a visiting scholar and guest lecturer at Oxford's Blavatnik School of Government, and Research Director of the £35 million UK and Australia funded Research on Improving Systems of Education (RISE) project. He has published over a hundred articles and papers (with more than 25 co-authors) and several books on a wide range of topics, including state capability, labor mobility, and education, among many others.

Nomthandazo Sithole advises the Ministry of Education in Liberia on the partnership schools initiative with the Tony Blair Institute for Global Change. She previously worked on youth and business growth initiatives in Liberia, her permanent home since 2011.

Contents

Introduction <i>by Lee Crawford</i>	1
15 years of education in Indonesia: rising enrolment and flat learning profiles – Amanda Beatty, Emilie Berkhout, Luhur Bima, Thomas Coen, Menno Pradhan, and Daniel Suryadarma <i>Commentary by Lant Pritchett</i>	5
Accountability and school performance: evidence from “Big Results Now” in Tanzania – Jacobus Cilliers, Isaac Mbiti, and Andrew Zeitlin	
Parental preferences and government delivery in Tanzanian primary schools – Andrew Zeitlin and Sam Solomon <i>Commentary by Aidan Eyakuze</i>	13
Teacher professional development around the world: the gap between evidence and practice – Anna Popova, David Evans, Mary E. Breeding, and Violeta Arancibia <i>Commentary by Celeste Carano and Nomtha Sithole</i>	19
Identifying the essential ingredients to literacy and numeracy improvement: teacher professional development and coaching, student textbooks, and structured teachers’ guides – Benjamin Piper, Stephanie Simmons Zuilkowski, Margaret Dubeck, Evelyn Jepkemei, and Simon J. King	

Scaling up successfully: lessons from Kenya's Tusome national literacy program

– Benjamin Piper, Joseph Destefano, Esther M. Kinyanjui, and Salome Ong'ele

Commentary by Barbara Bruns 25

Fair progress? Economic mobility across generations around the world

– Ambar Narayan, Roy Van der Weide, Alexandru Cojocaru, Christoph Lakner, Silvia Redaelli, Daniel Gerszon Mahler, Rakesh Gupta N. Ramasubbaiah, and Stefan Thewissen

Commentary by Jishnu Das 37

Familiarity does not breed contempt: generosity, discrimination and diversity in Delhi schools

– Gautam Rao

Commentary by Susannah Hares 47

Introduction

by Lee Crawford

This year's edition of the CfEE Annual Research Digest focuses exclusively on studies from across the developing world, including Africa, Asia, and South America. "Developing countries" (those below the \$12,000 GNI per capita threshold for high income countries) make up 36% of the world economy, but 83% of world population, and 87% of the world's school pupils. So as education economists we have every reason to be interested in their schooling experience.

There may also be more direct lessons for countries in the developed world. Though the specific challenges are inevitably very different when per student spending is measured in tens not thousands of pounds, there are principles that remain relevant. First and foremost is a focus on scale. In England, for example, one in three schools has now participated in a randomised trial, but most of these studies have involved fewer than 100 schools. Making the next step up to nationwide policy change is a whole different matter, requiring either experimentation at scale, or serious thinking about how systems function and the theory of scaling innovation. Fixing a dozen schools is very different to the thousands that we need to think about. This digest includes recent studies selected by leading thinkers on global education, and includes descriptive work highlighting the scale of the global learning crisis, causal studies investigating large-scale system reforms and new pedagogical approaches,

and going beyond test scores to look at the effects of schools on socialisation and social mobility.

The digest is opened by Lant Pritchett, Director of the largest ever British government-funded global education research programme (“RISE”). He discusses a descriptive paper highlighting both the scale and intractability of the global learning crisis. Indonesia is the fourth most populous country in the world. Over the last 15 years the country has made substantial efforts to improve the quality of its schooling, including a tripling of spending. This new paper demonstrates that despite these good faith efforts, progress has stalled, and at shockingly low learning levels. The lesson? We cannot be complacent; there are no easy answers; just doing more of the same is not enough; and we must look for new solutions. This story is broadly true across much of the developing world.

So what might work to improve things? Aidan Eyakuze (Director of leading African think tank “Twaweza”) reports on a pair of papers from Tanzania studying the government’s “Big Results Now” reforms, a major new high-stakes accountability exercise. A new national league table of schools based on standardised assessments led to improvements in test scores at the lowest ranking schools. This improvement was driven in part by strategic gaming of the measure – schools encouraged low performing students to drop out or not sit the test. But gaming did not explain all of it – there was real learning too. Encouragingly, a companion study focused on what parents actually care about, found that though they do highly value convenience, they also value learning.

To get better outcomes, schools and teachers need more than just accountability though; they need support too. Celeste Carano and Nomtha Sithole (Advisors with the Tony Blair Institute in the Liberia Ministry of Education) discuss a new paper from researchers at the World Bank

that establishes a new typology of teacher professional development programmes. Research shows huge variation in the effectiveness of teacher training programmes, so this paper sets out to document which features are most important for success (e.g. live lesson practice), and which are seen in large-scale government programmes (quite different features, as it turns out).

Barbara Bruns (Visiting Fellow at the Center for Global Development) discusses two papers on Kenya’s use of research evidence to design and scale-up a major reform of early grade literacy teaching. The design of the programme – Tusome – was based on an 800-school randomised trial that tested different combinations of programme elements. The most cost-effective combination included books for all children, coaching for teachers, and detailed teacher guides (“scripted lesson plans”). Keys to success were communicating clear outcome goals (hardwired into lesson plans) and supporting teachers in the classroom with regular visits from coaches.

Beyond test scores (which are of course important), education is also about shaping what kind of society we want. Schools are about more than just learning. We often talk in economics as if human capital was the prime motivation for governments to expand schooling, but other motivations abound. One might be social mobility. But are schools a leveller? Jishnu Das (Lead Economist at the World Bank) brings a note of scepticism. He focuses on a new World Bank report on social mobility. In the US, your chance of going to university is determined more by your ability at school than by your parent’s wealth. So if schools could be made to be fair, then equity in life chances should follow. Das shows that the same is not true in developing countries, raising questions about the ability of schools to level the playing field for later life chances. University attendance in Pakistan is better

predicted by parent's background than by test scores at school. The brightest poor kid is less likely to attend university than the least bright rich kid.

Finally, Susannah Hares (Global Education Director at the Center for Global Development) discusses Gautam Rao's Job Market Paper (recently published in the American Economic Review) looking at school desegregation in India. A policy providing subsidised places at private schools for poor children made the wealthy children nicer – more likely to share with and socialise with poorer classmates. Economists are usually sceptical of stated preferences: "talk is cheap". Where possible they like to look at revealed preference – what people actually do. So Rao devised a clever test, observing which kids were picked for a sports event: rich kids in integrated schools were indeed more likely to pick poorer kids.

Overall this is a fantastic set of papers from some of the top economists in the field. For those new to the field, I hope this digest helps to ease you into the world of global education economics, and to seasoned policymakers and practitioners, I hope you discover something new.

15 years of education in Indonesia: rising enrolment and flat learning profiles

– Amanda Beatty, Emilie Berkhout, Luhur Bima, Thomas Coen, Menno Pradhan, and Daniel Suryadarma

[Slide presentation](#)

Commentary by Lant Pritchett

Stagnant learning profiles in Indonesia: a warning against complacency

As research director of the RISE (Research on Improving Systems of Education) programme I see many great papers each year, but this year I would choose '15 years of Education in Indonesia: Rising Enrolment and Flat Learning Profiles' by Amanda Beatty, Emilie Berkhout, Luhur Bima, Thomas Coen, Menno Pradhan, and Daniel Suryadarma, from the RISE supported country research team in Indonesia. Let me not bury the lede.

The important fact is that the percentage answered correctly ("per cent correct") of simple multiple choice questions about arithmetic like $1/3 - 1/6 = ?$ (based on an IRT weighted index with each question adjusted for guessing) of a cohort of 18- to 24-year-olds in a (mostly) national representative household survey increased from 31.2 to 31.4% per cent in the 14 years from 2000 to 2014. That is an improvement of .2% per cent over 14 years, or the truly glacial pace of .017% per cent per year. Suppose Indonesia had the goal that its youth should be able to answer, say, even 50 per cent of eight simple

arithmetic questions. This would seem a necessary part of creating a generation of youth ready for the technological conditions and multiple challenges of the 21st century. At the current rate of change, achieving this modest goal would take over a thousand years $((.5-.314)/.017)=1112$). Indonesian youth would be ready for the 21st century in the 31st century.

A *descriptive* learning profile is just the bivariate relationship between some measure of learning (e.g. “ability to read”, “score on an assessment”, etc.) and grade attainment.¹ There are hundreds of estimates of grade attainment profiles – the website <http://iresearch.worldbank.org/edattain> contains grade attainment profiles of over 50 countries, with multiple periods for most countries, and which can be decomposed by sex, rural/urban, household assets. But, until recently there were few learning profiles as most assessments of learning, both international and national, are designed for a given grade (or age) and hence cannot compare the learning across many grades.

Recently there has been an expansion in learning profiles. One type are household surveys, pioneered by the ASER report of Pratham in India, that sample all children of a given village, both in and out of school, on a simple instrument to assess reading and numeracy. This approach, which has been replicated in a number of countries, provides a contemporaneous learning profile.

More recently, it has been discovered that surveys of adults with any measure of learning can be used to construct a retrospective descriptive learning profile for a current adult cohort. For instance, Pritchett and Sandefur (2017) use the Demographic and Health Survey literacy question to estimate

learning profiles for primary schools for 54 countries. Kaffenberger and Pritchett (2017) use the Financial Inclusion Insights data to estimate learning profiles for 10 countries.

This paper has two advantages over previous work. First, it uses the Indonesia Family Life Survey (IFLS) which is a household panel survey, representative of around 83 per cent of the national population, of about 30,000 individuals. The survey has been carried since 1993 with rounds in 2000 (Round 3), 2007 (Round 4), and 2014 (Round 5). This allows an estimate of the change in the learning profile over an extended period of 14 years. Secondly, there are eight numeracy questions in the IFLS with questions the curriculum expects to be mastered in grades 1 (two digit subtraction, e.g. $49-23=?$) to grade 5 level word problems (e.g. “If 65 per cent of citizens smoke, and the current population is 160 million, how many people do not smoke?”). This allows the authors to construct a more sophisticated learning measure than a single indicator of literacy (like the DHS) or the ASER-like indicators.

Figure 1 is the first bottom line of the paper: the learning profile is flat and falling even as grade attainment is rising.

In 2000 the likelihood a person with less than primary school education could answer a question correctly (adjusted for guessing as these are multiple choice questions) is 20 per cent. For people who have completed junior secondary this only rises to 30 per cent, and for those who have completed senior secondary (or progressed beyond that) this rises to only 40 per cent. This difference in capability between those with essentially no schooling and those who have completed secondary school is stunningly small. Assuming “less than primary” is two years of schooling and senior secondary is 12 years, this means that the per cent correct increases by just 2 percentage points per year of schooling completed.

¹ I am emphasising *descriptive*. There are no causal claims made here that the slope of the learning profile represents the causal impact on learning of an incremental year of schooling for a typical (or any) child.

This pattern in the IRT index is roughly true item by item. For instance, the percentage of people who can answer that $56/84$ is the same fraction as $2/3$ ($56/84=(2*2*2*7)/(2*2*3*7)=2/3$) only rises from 10% to 20% between “primary incomplete” and “senior secondary complete or more”. The flat learning profile is also true of those currently enrolled.

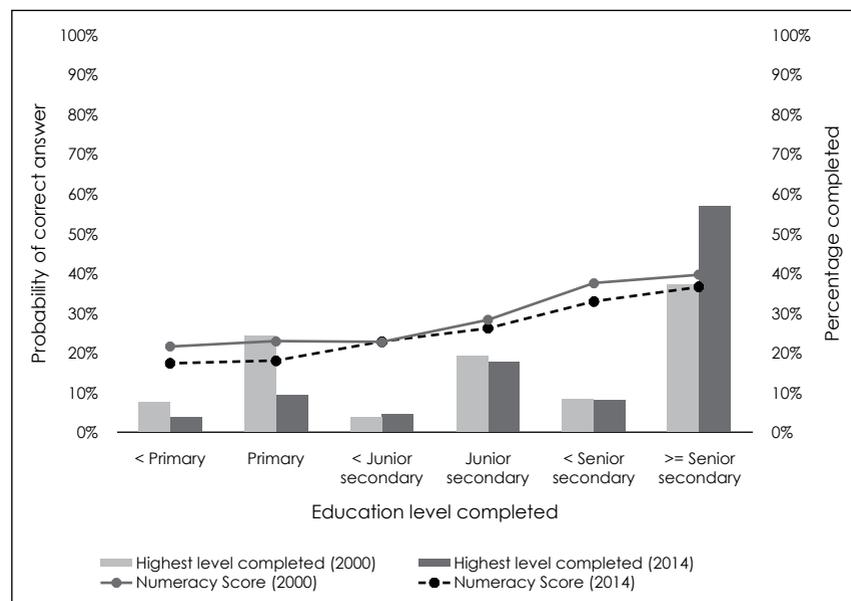


Figure 1: Numeracy learning profiles are flat and falling in Indonesia: 2000 to 2014

Source: Beatty, Berkhout, Bima, Coen, Pradham, and Surydarma 2018, figure 4.8.

But even more worrisome than that the learning profile is flat, it is falling. Those who were 18 to 24 years old in 2014 were less likely to answer questions correctly at each level of education completed. While the fall is small, the IFLS samples are huge so the fall is statistically significant.

Moreover, one would have hoped it were positive. You may recall that overall correct responses very slightly increased

over the period (from 31.2% to 31.4%); so how is it possible that the learning profile fell? An important feature of a learning profile of a cohort is that the level of measured capability/learning attainment can be exactly arithmetically decomposed into a grade (or level completed) attainment profile and a descriptive learning profile:

$$\begin{aligned} \text{Learning measure of a cohort} \\ &= \sum_{i=1}^N \text{Share of cohort with grade } i \text{ as highest grade} \\ &\quad * \text{Learning of those with grade } i \end{aligned}$$

This makes it clear that neither of the two components alone is sufficient for improvement in learning goals, and that the two interact. One could have a grade attainment goal (e.g. “every child finish grade 8”) but if the learning profile was flat (e.g. those with grade 8 complete did not have higher learning/competence/capability than those in grade 1), then grade attainment gain would not translate into a learning goal gain. Conversely, if one focused on the learning of those in grade 8, then one could increase that by holding back all of those in grade 7 who did not score well. By doing so, it would look like progress had been made on “learning in grade 8” but without any actual improvement having been made on the measure of cohort learning.

The second bottom line of the paper is displayed in Figure 2: the “per cent correct” of the 18- to 24-year-old cohort in 2014 was barely higher than that of 2000. The large gains in grade attainment of the 2014 cohort shown in Figure 1, where the fraction completing senior secondary or higher increased by nearly 20 percentage points, did not translate into large gains in the cohorts learned capability in mathematics because the learning profile was flat. Modest gains were offset by the deterioration in the learning profile, so that the gain was only 0.2 per cent. (Even had the learning profile stayed constant the increase would have only been 3.5 per cent.)

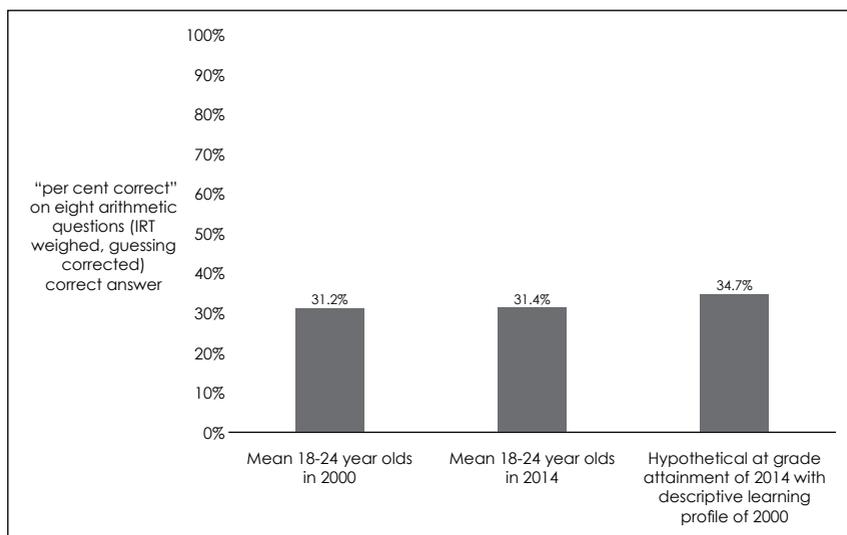


Figure 2: The numeracy capability of Indonesia's youth barely increased at all from 2000 to 2014, in spite of gains in grade attainment, as the learning per grade fell

Source: Calculations of the authors of the paper

The reason I think this paper is the most important paper this year in the economics of education in developing countries is that I think the biggest danger in the domain of education is complacency, of two types, and this paper challenges complacency of both types precisely because the paper is about Indonesia.

The first type of complacency is that learning performance will improve "naturally" or "inevitably" over time, or as a consequence of other broad positive changes in the economy, polity, or society. If this were true of anywhere, it should be true in Indonesia where there has been rapid progress between 2000 and 2014 on many fronts. Indonesia has made an impressively smooth transition from a long period of authoritarian rule to a stable, competitive democracy following the resignation of the long-term

president, Suharto, in May of 1998. It is one of the few countries in the world where the indicators of government capability over the period 1996 to 2012 show improvement (Andrews, Pritchett, Woolcock 2016). GDP per capita more than doubled over this period. A quite radical decentralisation of government responsibilities and budget to districts, intended to improve governance, was undertaken in 2000. So this (mild) deterioration in learning performance did not happen in a failing state, or even a stagnating state, but in a polity and economy that were, in many ways, looking positive.

The second type of complacency is that "we" (global and national education experts) know what to do. Once learning is prioritised, it will be easy to make progress with "business as usual" intensified, or, crudely put, BAUWMM (business as usual with more money). Again, Indonesia shakes that complacency. Indonesia has a strong commitment to spending on education and has legal commitments to devote 20 per cent of revenues to education. This means education spending has tripled in Indonesia. Part of this was allocated to a doubling of teacher pay which was intended to raise the performance of teachers, and thereby standards. But a rigorous study of the impact of this policy change, entitled 'Double for Nothing' (de Ree et al 2017), found the results very much to the contrary.

That a country making progress on so many fronts; that has undertaken major increases in education spending; that has pursued many standard programmes and policies to improve education; and that succeeded in expanding grade attainment (increasing the fraction completing secondary school by 20 percentage points), has made very near zero progress in 14 years (albeit on one measure of learning) should shake everyone out of complacency.

References

- Andrews, M., Pritchett, L., and Woolcock, M. (2016), 'The Big Stuck in State Capability for Policy Implementation'. Center for International Development Working Paper No. 318. Harvard University.
- Beatty, A., Berkhout, E., Bima, L., Coen, T., Pradhan, M., and D. Suryadarma (2018). '15 years of education in Indonesia: rising enrolment and flat learning profiles.' Paper Presented at RISE Conference July 2018.
- de Ree, Joppe; Muralidharan, Karthik; Pradhan, Menno; Rogers, Halsey. (2017). 'Double for Nothing?: Experimental Evidence on an Unconditional Teacher Salary Increase in Indonesia.' Policy Research Working Paper No. 8264. World Bank, Washington, DC. © World Bank. <https://openknowledge.worldbank.org/handle/10986/28987> License: CC BY 3.0 IGO.
- Pritchett, L. and J. Sandefur. (2017). 'Girls' Schooling and Women's Literacy: Schooling Targets Alone Won't Reach Learning Goals'. RISE Working Paper 17/011 (February 2017).

Accountability and school performance: evidence from "Big Results Now" in Tanzania

– Jacobus Cilliers, Isaac Mbiti, and Andrew Zeitlin

[Slide presentation](#)

Parental preferences and government delivery in Tanzanian primary schools

– Andrew Zeitlin and Sam Solomon

Unpublished

Commentary by Aidan Eyakuze

Two themes have recently dominated the public debate on education in Tanzania. The first was prompted by a **steep decline** in the rankings of historically high-performing public schools in Tanzania (**ACSEE 2018**). In 2018, many of the schools which in the past had ranked near the top in terms of exam results were languishing towards the **bottom of the national league tables**. **One girls' school** in particular came in for some deeply critical, almost insulting criticism. What happened to these schools? Can their disastrous performance be reversed? If so, how?

The second was the issue of parental contributions to improving the quality of primary education. In the context of a well-publicised government policy of fee-free basic education, any suggestion that parents might contribute

towards improving school infrastructure (or even contributing to the cost of school meals) were met with **strong pushback** from both cash-strapped parents and vote-conscious politicians. Headteachers faced summary **dismissal** for even calling meetings to discuss parental contributions. Ministerial **clarifications** on this issue – school administrators can solicit but not compel voluntary parental contributions and must obtain permission from local government authorities to do so – were timid. The result: more **confusion**.

These two papers speak directly to these two pressing issues. And they both suggest that the solution lies in providing salient information to those who take the decisions that make the biggest difference to children’s learning outcomes: school headteachers and parents.

The first paper by Jacobus Cilliers, Isaac Mbiti and Andrew Zeitlin (2018) uses evidence from a major Tanzania education reform programme called “Big Results Now” (BRN) to answer the following question:

‘Can public rankings improve school performance?’

BRN was a systemic, top-down accountability reform that published both nationwide and within-district school rankings, among other policies, starting in 2013. Combining data on the universe of school performance from 2011–2016 with EMIS and Service Delivery Indicator datasets from the post-reform period, we identify the impacts of the dissemination of within-district rankings by using a difference-in-difference strategy that exploits a key feature of such rankings—that they exert pressure in particular on schools at the top and bottom of their respective districts. We find that BRN improved learning outcomes for schools in the bottom fifth of their districts. While we find no evidence that these

gains resulted from zero-sum transfers of resources from schools at the center of their district distribution to schools at the tails, we also find no evidence of impacts on physical, human, or financial inputs into schools. By contrast, we do find that the number of test-takers falls in schools facing pressure from within-district rankings; this appears to be driven by strategic exclusion of students rather than forced repetition within schools or sorting across schools.

I like this paper for two main reasons. First, the authors did not spend copious amounts of money and time (which is not to say they skimmed on the level of effort) to collect new data. They used existing administrative data and leveraged the World Bank’s investment in producing service delivery indicators (SDIs). Their substantial value addition was in applying the powerful difference-in-difference methodology to surface their findings and conclusions.

The second reason I like it is because the solution hit me between the eyes as soon as I saw the presentation: just show headteachers how their school ranks relative to others in the district, and watch them respond. Simple and cheap. Information on district school rankings does matter for learning outcomes.

However, there is a downside. If schools can exclude weaker students from taking the final exams in order to boost rankings, some will probably do so. Encouragingly, the study finds that overall, despite this “gaming”, there is still a real net positive improvement from the new incentives. The question for policymakers is whether they can prevent the negative consequences.

If the first paper seeks to understand the effects of salient information on school managers, the second examines

matters from the consumer's perspective. Parents' views are important, but, they are often ignored. They are rarely asked what they think or care about.

The second paper entitled 'Parental preferences and government delivery in Tanzanian primary schools', by Andrew Zeitlin and Sam Solomon, asks:

'What do Tanzanian citizens value in their primary education system?'

Zeitlin and Solomon hypothesise that knowing what parents want will help governments do a better job of delivering value, while encouraging parents to engage with a school system that is responsive to their expressed needs.

Using Africa's largest mobile phone survey platform, "Sauti za Wananchi", they ask parents 'to express their preferences between alternative pairs of [primary] schools – to understand the extent to which parents value proximity, infrastructure, teacher-pupil ratios, and learning outcomes, in the primary schools to which they have access.'

They find three things. First, parents prefer schools that are close and that produce learning. Secondly, this preference varies across regions (i.e. it is neither static nor universal across the country). The paper notes for example that 'households in Mara region (north west Tanzania), for example, reveal a willingness to walk 1.86 km for an improvement of 10 points in average exam scores while respondents in Pwani region (on the coast adjacent to Dar es Salaam) reveal they are only willing to walk 0.64 km for the same improvement.'

Finally, they find that there is a gap 'between the strength of parents' preference for learning over convenience and

the efficiency with which regions actually deliver learning outcomes, suggesting scope for improved alignment between government and citizen priorities.'

This paper tells us what parents want from primary schools, and with what intensity, and it also highlights how service delivery results fall short of their expectations. These have hitherto been understood only through anecdote. We now have a stronger evidence base – client preferences – from which to make decisions on where schools are built and how they are resourced. Will Tanzania's education administrators have the courage to incorporate such feedback in their decisions? I hope so.

References

- Cilliers, J., Mbiti, I., and A. Zeitlin. (2018). 'Accountability and school performance: evidence from "Big Results Now" in Tanzania.' Paper Presented at RISE Conference July 2018.
- Zeitlin, A. and S. Solomon. (n.d.). 'Parental preferences and government delivery in Tanzanian primary schools.' Unpublished paper.

Teacher professional development around the world: the gap between evidence and practice

– Anna Popova, David Evans, Mary E. Breeding, and Violeta Arancibia

[Working paper version \(free\)](#)

Commentary by Celeste Carano and Nomtha Sithole

Governments and donors invest in teacher training with the expectation that it will produce better teachers and higher student achievement. But is that spending effective? In countries with low resources, like Liberia, where the government has less than \$50 per child for education, up to \$5000 may be spent training a teacher. That's a high cost for a programme without a sure outcome. Policymakers and donors can also choose from a variety of approaches, ranging from mass training programmes, to peer-to-peer support, to coaches, to tech-based resource libraries. But what will best support teachers and results in the classroom?

In this paper, the authors ambitiously aim to tackle both the question of what characteristics define successful teacher training programmes and how to define them. They set out to identify what aspects of teacher training programmes are linked with better student learning outcomes through a tool to measure and compare teacher professional development programmes across countries. The In-Service Teacher Training Instrument (ITTSI) aims to capture the details of teacher training programmes and is piloted in this paper.

The ITTSI is wide-reaching, capturing indicators including design and implementation of programme, scale, cost, who is targeted, what knowledge and skills it intends to develop in teachers, how the programme is delivered to teachers and whether it was positively or negatively received by participants. The authors first use the tool to analyse 33 programmes with existing impact evaluations, sorting out those that were effective from those that were not. Importantly, after coding information from the papers, the authors made considerable effort to obtain additional data not shared in the papers, by contacting authors and implementers directly, ensuring a high proportion of the ITTSI was completed for each programme. The authors then analysed which aspects of the professional development programmes were associated with the largest improvements in student learning.

Building on that both quantitative and qualitative analysis, they worked with World Bank regional teams to identify countries for phase two of the analysis of at-scale programmes, ultimately collecting 48 ITTSI surveys of at-scale programmes, in addition to 91 short versions of the survey (the aptly named 'BITTSI').

Ranking the 33 previously evaluated programmes by their standardised impact on student test scores, the researchers selected the top 16 of the programmes as the 'best'. They then compared the average value of the 'top performer' programme indicators to those of the at-scale, government-led programmes to investigate the difference in practices between the successful professional development programmes and government programmes.

So what are the features of training programmes that proved successful in impact evaluations? Most notably, programmes which took an intensive, practical 'hands-on' approach appeared as standouts. Teacher practice in the classroom,

consecutive days of face-to-face training, and follow-up visits to review material from the training, were all associated with gains in student performance. In qualitative interviews with programme implementers, the researchers found that teachers most enjoyed the interactive, fun programmes that enabled more active learning. Perhaps unsurprisingly, more active learning for teachers seems to mean more active learning for their students as well. Other successful features included linking career opportunities to training, and targeting training to teachers based on their years of service (although only 2 of the 33 programmes took these approaches).

Government programmes, in contrast, were less likely to be linked to career opportunities, suggesting they lack incentives for teachers to improve after the training. In addition, fewer at-scale government programmes provide reading materials and books – linked with student learning gains – than do the 16 high-performing programmes. Most striking, the hands-on and in-person aspects of successful programmes were all less likely to be found in the government programmes, or found to a lesser degree, meaning teachers had fewer days of face-to-face training, fewer follow-ups from the trainers, and less time to practice what they learned first-hand in an applied way.

Readers may come away questioning exactly why the government programmes differ in these ways, and cost comes to mind immediately as a potential factor. But while the authors gathered information on programme cost with ITTSI, analysis of this was not shared in the paper. This leaves an open question of how the 'best' professional development programmes and the at-scale, government implemented programmes compared on the cost per teacher, and overall cost-effectiveness. While the readers, and authors, assume that at-scale programmes are lower cost, the size of the gap would be interesting to explore further and particularly

relevant for policymakers considering trying to replicate more (presumably) expensive aspects of the successful programmes. While in-person follow up and continued practice with peers may translate to greater student gains, governments would need to consider not only the higher costs of adding that more staff-intensive time into training, but also just what human resources can be allocated to enable that follow-up – likely at an additional financial cost.

The differences in intent and follow-through in the design of government programmes is also touched on, but only briefly. The authors do acknowledge that due to the scope of the study, their inability to track programmes from design through to classroom practice make it difficult to identify impact and the point at which the weaker programmes fail to deliver. Our own observations are that professional development programmes may tend to be developed in reaction to pressures to improve school results – in the hopes they will produce a quick fix – and are often slow to evolve to actual needs. In Liberia, an emergency primary school teacher training programme developed as a temporary solution after the war remains the default training programme fifteen years later. With the slow pace of innovation and adaptation to evolving education systems, developing countries continue to channel money to programmes like this with low or no returns on student learning outcomes.

Similarly, the balance of politics is a challenge when considering how to integrate teacher incentives into professional development programmes. This may be why nearly half of all government programmes in the study did not include this feature. The role of unions in policy and in demanding incentives for teachers often enhances tensions in decisions around the use of resources. Professional development, however effective or ineffective, is popular;

programmes that measure, monitor, and require teachers to meet explicit standards or face consequences may be received with less enthusiasm.

But this only makes this paper even more valuable for policymakers (and donors). The most effective programmes have qualities that may (we presume) make them more expensive, more time consuming, and more difficult to deliver. But with some evidence that these qualities translate into impact for students, perhaps the argument for funding can be more easily made. In addition, ITTSI alone holds value to policymakers and programme managers and will become more valuable the more it is applied to future teacher training programmes. If used at the World Bank alone, that would build up a stronger comparison base for other programmes in the future. The ITTSI could ultimately help strengthen the impact of professional development programmes by probing their design and anticipating points of impact prior to roll-out. While to date a relatively small number of programmes have been evaluated using ITTSI, as the authors rightly point out, a key purpose of the study was to design and test the tool for further use in the future – increasing its value over time.

References

- Popova, A., Evans, D., Breeding, M.E., Arancibia, V. (2018). 'Teacher Professional Development around the World: The Gap between Evidence and Practice (English).' Policy Research working paper no. WPS 8572. © World Bank Group, Washington D.C.

Identifying the essential ingredients to literacy and numeracy improvement: teacher professional development and coaching, student textbooks, and structured teachers' guides

– Benjamin Piper, Stephanie Simmons Zuilkowski, Margaret Dubeck, Evelyn Jepkemei, and Simon J. King

[Published version](#)

Scaling up successfully: lessons from Kenya's Tusome national literacy program

– Benjamin Piper, Joseph Destefano, Esther M. Kinyanjui, and Salome Ong'ele

[Published version](#)

Commentary by Barbara Bruns

Kenya's reading revolution

Three years ago, the world woke up to the fact that “schooling ain't learning” and set education SDGs (Sustainable Development Goals) that for the first time focused on what children actually learn in school. Since then, the news has been all bad. In Indonesia, the average math skills of adolescents have not improved in 14 years (as Pritchett notes earlier in this volume). In rural India, after three years in school, 68 per cent of students cannot read a word of English

(ASER 2007). In Chad and Niger, over 85 per cent of children finishing primary school cannot read and understand a text (PASEC 2014). Across Africa, the share of grade 2 students who cannot not read a single word is 64% in Uganda (English), 56% in Zambia (Chitonga), and 90% in Malawi (Chichewa) (USAID Early Grade Reading Barometer). UNESCO's official estimate is that 56% of all primary school-age children worldwide are not achieving basic literacy, and in two-thirds of cases this is despite actually completing primary school.

In this depressing context, there is important news coming out of Kenya. Two new publications by Ben Piper and co-authors document big improvements in early grade reading over the past few years and their research findings on how Kenya has done it. The first paper provides the most systematic evidence to date on how to improve teachers' ability to teach reading in the first two grades of primary school. The second paper provides evidence on how to take pilot programmes to national scale. This is a question of huge policy relevance but near-zero research base. Together, the two papers are an object lesson in how to do research with real-world impact.

Kenya's education ministry deserves kudos for both research advances. Since 2012, the government has engaged in an all-out effort to transform the way basic reading and numeracy skills are taught. Importantly, part of the strategy is serious evaluation of these efforts. The first stage was several randomised evaluations of a pilot Primary Mathematics and Reading (PRIMR) programme. The first of the papers I reference here, by Benjamin Piper, Stephanie Simmons Zuilkowski, Margaret Dubeck, Evelyn Jepkemei, and Simon J. King, for *World Development* journal (2018), reports on one of these – an evaluation of PRIMR in a sample of 834 schools, a large enough sample to compare three different programme designs with a control group of schools.

Next, the ministry used evaluation results to design a new national literacy programme, "Tusome" (Swahili for "Let's Read"). Finally, and perhaps most unusually, in the national scale-up of Tusome the government opened all levels of the education system to classroom observation, data collection and external analysis of the quality of the implementation process.

So, what works to help teachers impart literacy?

The past 15 years have seen a big increase in experimental evidence on improving teachers' effectiveness. Evidence supports matching the curriculum to the pace of students' mastery ("teaching at the right level"), remedial education (sometimes using computers) for students falling behind, providing a book to every child, some forms of teacher training, individualised coaching of teachers, and, somewhat controversially, scripted lesson plans which guide teachers through each day of the curriculum.

Several meta-studies have synthesised this research, but limitations in the underlying studies make it hard to draw clear conclusions. Few studies test alternative programme designs, so there is little systematic evidence about different programme elements. Studies typically evaluate one intervention design in a single country, so it's not always clear things would work in a different setting. Finally, many evaluations fail to analyse costs, making it impossible to generate the evidence policymakers need most – not just on the effects of different interventions, but their comparative cost-effectiveness.

The PRIMR programme impact evaluation was designed to give the Kenyan ministry clear answers to these questions:

1. Which programme elements are most essential?
2. Are there important complementarities among elements? And, if so:
3. Which combination is most cost-effective?

PRIMR was designed in a context where: less than 5% of first- and second-grade children met the government literacy benchmarks; 80% of teachers reported no professional development support during the prior year; and a network of curriculum support officers existed, but lacked the time and transport to visit classrooms.

The government devised three different strategies for improving reading instruction and tested them in a systematic way. Each variant systematically added intensity (and costs) to the prior one, so a randomised trial would reveal whether the additional ingredients (and costs) were really worth it.

Strategy 1: Teachers each received ten days per year of professional development. Curriculum support officers received 15 days of training on tablet-based teacher observation, feedback, and coaching tools. Training and support were focused on the existing curriculum and available reading materials (where less than half of students typically had textbooks).

Strategy 2: Teachers received the same training and coaching time, but were given an additional new set of textbooks. These new books were based on the latest research on how to teach literacy and were provided to students on a one to one ratio. Teacher and curriculum support officer training were focused on using the new books and teaching techniques. Teachers were encouraged to develop their own lesson plans for using the new books.

Strategy 3: Teachers received the same training and coaching time and new textbooks as in Strategy 2, with the addition of new teachers' guides matched to the new textbooks. The guides contained 150 days of partially scripted lesson plans.

The three-level PRIMR evaluation was implemented in grade 1 and 2 classrooms in 847 government schools in rural Kenya between March 2013 and October 2014. As curriculum support officers are organised by zone, 44 schooling zones were randomly assigned to one of the three treatments or the control group. Baseline and endline reading skills against the government benchmarks were measured using the EGRA (Early Grade Reading Assessment), which generates scores on letter recognition, oral reading fluency, and reading comprehension.

Key results

As shown in the chart below, the study found very limited impact from professional development and coaching using existing materials (treatment 1). Adding the new books led to statistically significant improvements in English and Kiswahili (treatment 2). Adding the teacher guides led to very large improvements in student literacy (treatment 3). The average effect size of RCTs in Africa reported by Conn (2017) is of the order of 0.23 SD). PRIMR produced results above 1 standard deviation. There are some reasons to quibble with these results. The evaluation had to rely on a difference-in-differences estimation overlaid on the randomisation because the treatment groups were not perfectly balanced at baseline. But it is highly unlikely that effects this large or consistent are a fluke.

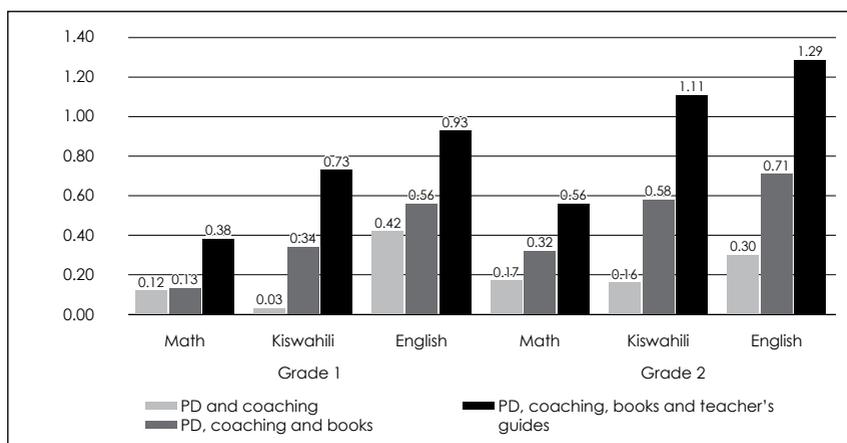


Figure 1: PRIMR evaluation: key results

Source: Piper, Zuilkowski, Dubeck, Jepkemei, and King (2018), Figure 1.

Value for money?

It is great that the intensified treatments produced stronger impacts, but what was their marginal cost? The cost of the core treatment (training and coaching) was kept constant across the three variants at US\$5.63 per pupil. The addition of revised textbooks for all students cost an extra US\$2.38 per pupil. The teachers' guides added just US\$0.16 per pupil. Although it's unclear whether all the costs of content development were included, the guides were a pretty low-cost addition.

Piper and co-authors estimate the number of students able to reach the Government's oral reading fluency benchmarks for every additional \$100 spent on each treatment.

Treatment 1 (training and coaching only): 2 extra students able to meet the oral reading fluency benchmarks (65 words per minute for English, 45 for Kiswahili)

Treatment 2 (training + books): 6 to 8 more students reaching the benchmarks.

Treatment 3 (training + books + guides): 15 more students reaching the benchmarks.

The large learning gains produced by the teachers' guides doubled the cost-effectiveness of the programme.

Because of its careful design, the PRIMR impact evaluation has made an outsized contribution to our understanding of how different elements combine to raise the effectiveness of teachers. This study demonstrates the payoff to randomised trials that explore the marginal impact, and costs, of complementary programme elements. This is more expensive research – three treatment arms require a sample of 800+ schools rather than the 100-200 school sample seen in most RCTs. But if the knowledge yield from this single study influences other governments and research teams to follow suit, we can expect bigger dividends from future research on education quality.

What does it take to scale up a successful pilot programme?

Given PRIMR's positive results, in 2015 the Kenyan government decided to scale up the third variant of the programme to all 23,000 public primary schools (and 1500 low-cost private schools). After two years, an independent external evaluation documented Tusome's sizeable impacts on students' reading fluency (Freudenberger and Davis 2017).

The second of the papers I reference here, by Benjamin Piper, Joseph Destefano, Esther M. Kinyanjui, and Salome Ong'ele for the *Journal of Educational Change* (2018), is unusual in that it focuses on the implementation processes underlying the

results. This topic is central to government policy but rarely gets research attention.

The authors compare Tusome to the conceptual framework developed by Crouch and de Stefano (2017) on why it's so hard to drive system-wide change in education. Crouch and de Stefano observe that system-level change requires getting decentralised schools and teachers to adopt new behaviors. They posit that an education ministry's ability to do this rests on its institutional capacity for three key functions:

1. setting and communicating expectations for the outcomes of education;
2. monitoring and holding schools accountable for meeting those expectations; and
3. intervening to support the students and schools that don't meet expectations.

Piper and coauthors document how Tusome's design addressed, and strengthened, some of these core functions.

First, on expectations, the programme is organised around two clear outcome goals: national benchmarks for oral reading fluency (words per minute) in Kiswahili and English. Piper and coauthors find these have been communicated broadly and extensively through training and coaching programmes. These outcome goals are also hardwired in the design of the textbooks and teacher guides, which present a sequence of lessons geared towards achieving "emergent" and "fluent" reading at appropriate points. Classroom observations in 2017 found that 99% of classrooms had one book per student, and 95% of teachers were using the guides. Moreover, teachers reported that they believe the guides help their students' progress.

Second, school monitoring is a key element of Tusome. Curriculum support officers make regular classroom visits using tablets with instructional support tools. They are reimbursed for travel to schools, where they spot check students' reading and observe teaching. They record where teachers are in the curriculum and whether teachers are using the new techniques. They meet with teachers to provide one-on-one feedback. In 2016, curriculum support officers averaged 90 visits each and recorded and uploaded 113,604 classroom observations. More than 80% of Class 1 and 2 teachers reported being observed by a support officer at least once per term. While this was short of the targeted three times per term, it is a degree of classroom-level monitoring and data collection that is unprecedented in Kenya, and rarely seen anywhere.

Third, on targeted support, Piper and coauthors find this was the weakest part of the implementation. While the school system is for the first time generating real-time data that clearly expose variations in performance at the school and district level, the researchers document little action thus far to target resources or support interventions to those that are struggling.

The 2017 external evaluation measured student reading in a national sample of schools after two years of implementation. In both English and Swahili, in both Class 1 and Class 2, there were large gains on a wide range of reading tasks. The percentage of Class 2 children meeting the national benchmark approximately doubled for both English (34% to 65%) and Swahili (37% to 66%). These are impressive gains for a programme scaled up nationally in just two years.

Piper and co-authors believe the main driver of success is the Ministry of Education's effectiveness in the first two core functions identified by Crouch and De Stefano (2017)

– setting and communicating expectations and monitoring implementation. The national programme achieved high implementation fidelity in materials provision, teachers’ professional development and, to some extent, instructional support.

More profoundly, they believe Tusome has transformed the “instructional core” of the first years of schooling. The programme moves teachers to engage with their students in a new way, with new teaching techniques, new materials, and new expectations for learning outcomes.

They also note that Tusome has reoriented the education system to focus on the classroom. Regular visits, structured observations, and feedback replace the isolation and performance vacuum found in most school systems at the classroom level. While there are no incentives or sanctions for teachers associated with the observations and feedback, the simple fact of classroom-level monitoring has reshaped the system’s norms and teachers’ felt accountability for performance.

Tusome is still only a few years old, and it will be important to watch how it is sustained, deepens, or fades, over time. Big improvements in early grade reading should produce lower grade repetition and better learning outcomes as children move up through primary school; administrative data will help confirm that Tusome’s early effects are sustained. But through careful research on what works to improve reading, and how programmes can be scaled successfully, Kenya and an enterprising group of researchers have made big contributions to our knowledge base. Continuing evidence from Kenya hopefully will expand that base further – and inspire other countries and research teams to do similarly important work.

References

- ASER Centre (2017), Annual Status of Education Report (Rural) 2016, New Delhi, http://img.asercentre.org/docs/Publications/ASER%20Reports/ASER%202016/aser_2016.pdf
- Conn, K.M. (2017). ‘Identifying effective education interventions in sub-Saharan Africa: A meta-analysis of impact evaluations.’ *Review of Educational Research*, 87(5), 863-898.
- Crouch, L., and J. de Stefano. (2017). ‘Doing Reform Differently: Combining Rigor and Practicality in Implementation and Evaluation of System Reforms.’ International Development Working Paper No. 2017-01, RTI International.
- Freudenberger, E., and J. Davis. (2017). ‘Tusome External Evaluation – Midline Report.’ https://pdf.usaid.gov/pdf_docs/PA00MS6J.pdf
- PASEC (2015). ‘PASEC2014 Education System Performance in Francophone Sub-Saharan Africa: Competencies and learning factors in primary education.’ *Conférence des ministres de l’Éducation des États et gouvernements de la Francophonie, Dakar* http://www.pasec.confemen.org/wp-content/uploads/2015/12/Rapport_Pasec2014_GB_webv2.pdf
- Piper, B., de Stefano, J., Kinyanjui, E.M., and S. Ong’ele. (2018). ‘Scaling up successfully: Lessons from Kenya’s Tusome national literacy program.’ *Journal of Educational Change*, 19(3), 293-321.
- Piper, B., Zuilkowski, S.S., Dubeck, M., Jepkemei, E., and S.J. King. (2018). ‘Identifying the essential ingredients to literacy and numeracy improvement: Teacher professional development and coaching, student textbooks, and structured teachers’ guides.’ *World Development*, 106, 324-336.

Fair progress? Economic mobility across generations around the world

– Ambar Narayan, Roy Van der Weide, Alexandru Cojocaru, Christoph Lakner, Silvia Redaelli, Daniel Gerszon Mahler, Rakesh Gupta N. Ramasubbaiah, and Stefan Thewissen

[Published version \(free\)](#)

Commentary by Jishnu Das

Should I stop kidding myself that learning leads to educational opportunity?

We want to make sure that children learn. We agree that schools aren't delivering. But will policies to improve children's test scores also enhance equity and opportunity in an increasingly unequal world?

I worry that they won't even come close.

Think about it this way. Learning is a goal in itself; it is hard to participate with dignity in a globalised world without knowing how to read or write. But that's not the only goal we are selling; we are also arguing that "learning is a ladder to opportunity"—what holds the poor back is the low quality of schools they now attend in increasingly higher numbers.

And in case you have any doubt that the poor are indeed held back, the World Bank's recent report (2018) on education mobility makes for sober reading:

Mobility from the bottom half of the education ladder to the top quartile has fallen over time in developing economies, whereas persistence at the bottom has increased. In the median developing economy for the 1980s generation, less than 15 per cent born into the bottom half make it to the top quarter, while more than two-thirds stay in the bottom half [...] These concerns are heightened for Africa and South Asia, where the prospects of children are still tied to the socioeconomic status of their parents more closely than in any other developing region, which suggests that relative mobility in these two regions will continue to be low in the near future.

In the U.S., test scores level the playing field ...

One explanation for low mobility is that children from disadvantaged backgrounds are not learning in school and this prevents them from accessing the higher levels of education that children from more privileged backgrounds regularly reach. The lack of mobility in low-income countries tracks back to differences in test scores so that fixing schools will help build a fairer society.

This is indeed what we see in the U.S. Carneiro and Heckman (2003) showed that once you control for test scores, the difference in college attendance by socioeconomic status vanishes. Rather than the picture on the left, where academic ability has little to do with years of schooling, the U.S. looks like the picture on the right. Those from disadvantaged backgrounds still get fewer years of schooling, but these differences can be fully traced back to their academic performance. This powerful finding immediately puts school quality at the centre of all debates about social mobility; in the U.S., learning is indeed a (maybe the) ladder to opportunity.

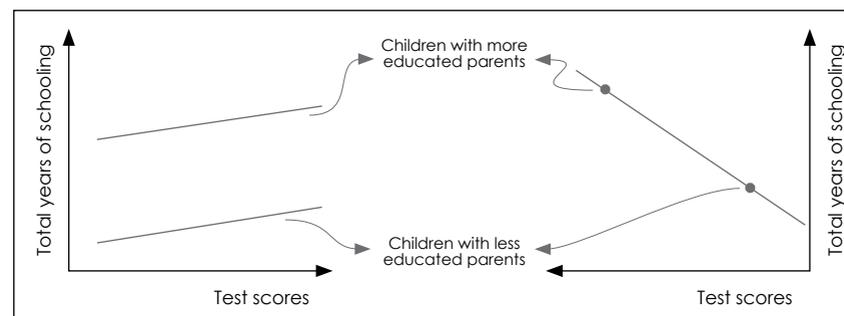


Figure 1: Societies can have high returns to family background (left) or high returns to academic achievement (right)

...but in low-income countries, not so much

Similar data has been hard to come by in low-income countries because we need to know how test scores earlier in life lead to (more) schooling later, which requires long follow-up studies. But as these studies start to emerge, the picture looks very different. With Andres Yi Chang, I have been looking at new data from the Learning and Educational Achievement in Punjab Schools (LEAPS) project, which has now followed 5,000+ children from ages 7 to 9 through to their twenties. For around 700 of these children we have test scores from their primary schooling years.

Figure 2a and 2b plot child test-scores from 2003 to 2011. In 2003, the children were in Grade 3 and you can roughly think of each year as the subsequent grade, so that in 2006 most of these children are in Grade 6 (with the exception of those that are held back). The two lines are for children from “high” and “low” socioeconomic backgrounds, defined as an index of parental education and household wealth. Two basic facts emerge from this picture.

- By the time they are in Grade 3, there is a difference in the test scores of high and low socioeconomic

(SES) children. This gap corresponds roughly to how much children learn in a year, so children from low SES backgrounds are one year behind children from high SES backgrounds by the time they are in Grade 3. Through the primary years, there is a slight divergence, but this is neither statistically nor qualitatively significant.

- By 2011, when (most of) these children are 17, large gaps of one standard deviation have emerged. As Figure 2b shows, this widening gap is entirely due to differential drop-outs. Among children who are still in school, the gaps remain the same as at the end of primary school. But because 61% of children from low SES backgrounds drop out of school compared to 35% from high SES backgrounds (we track and continue to test those who drop out), the average gaps in the cohort are much wider.

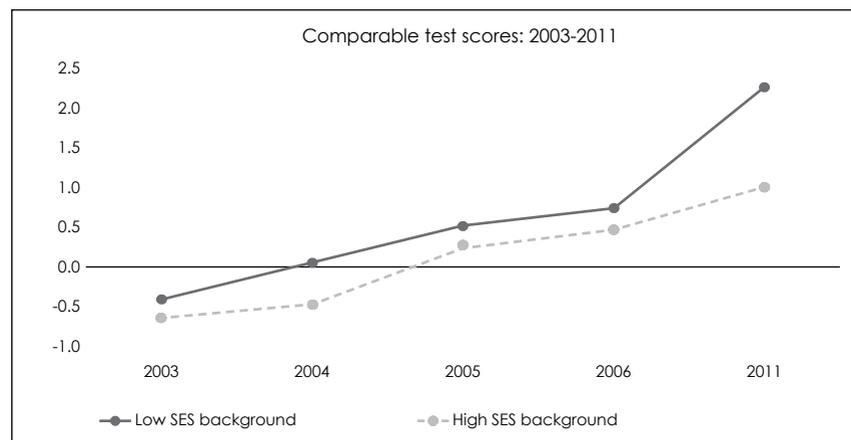


Figure 2a: Test score gaps have developed by Grade 3, remain steady over the primary schooling years, and then widen dramatically ...

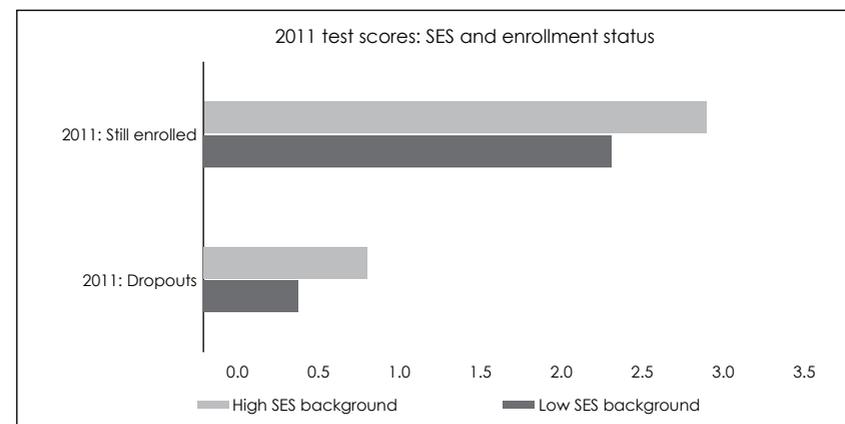


Figure 2b: ... because children from low SES backgrounds drop out of school at higher rates than those from high SES backgrounds

Note: These graphs are based on longitudinal data collected as part of the LEAPS long-term follow-up study and are taken from ongoing work by Das and others on educational mobility and test scores in low-income countries.

Figure 3 then plots the number of years of completed schooling in 2017 on the vertical axis (14 years after we first surveyed and tested these children), and the test scores of the children in 2006 on the horizontal axis. We show this relationship between test scores and completed years of schooling for children whose fathers have no education versus those whose fathers have more than middle-school education.

- Although higher test scores are correlated with more years of schooling, family background is paramount. The highest performing children whose fathers have no education complete just as many years of schooling as the lowest performing children whose fathers have more than middle school education (9–10 years).

- In regressions, every additional year of maternal and paternal education is associated with 0.42 more year of schooling, so that five years of maternal and paternal education implies two more years of schooling for a child. To get the same boost in years of schooling from higher test scores in third grade would require a two standard-deviation increase, equivalent to moving a child from the 50th to the 97th percentile of the test score distribution.

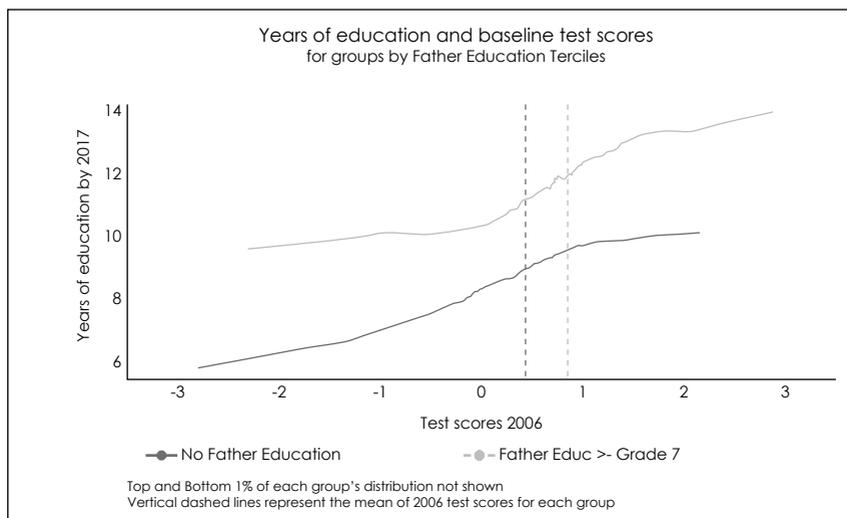


Figure 3: Children with higher test scores go on to complete more years of schooling ... but family background remains critical.

Note: This figure is based on data from the LEAPS long-term longitudinal study.

Unfortunately, we don't have a single intervention that has been shown to boost learning by even one standard deviation over the primary schooling years. In fact, based on these estimates, here is what we could honestly say to an illiterate

parent wondering what our latest learning intervention will do for her child:

Our intervention boosts learning by 0.5 standard deviations, which will then increase your child's years of schooling by 6 months. But even if this boosts your child to the 99th percentile of the test score distribution, she will still get less education than the children of parents with more education, even if they are at the 1st percentile of the test score distribution. But this is the best we have.

Fixing schools and improving learning will only go a small way towards addressing the deep equity issues we face today.

Yes, but, what if ...

We can quibble. Maybe these results are specific to Pakistan. But work we are now doing with Abhijeet Singh finds remarkably similar patterns in the Young Lives countries of Peru, Ethiopia, Vietnam, and India. Maybe test scores have an independent effect on wages and later life outcomes regardless of their association with further educational attainment. This is again true for the U.S., where wages become more correlated with test scores and less correlated with educational attainment as workers gain labour market experience. But emerging evidence from China, Brazil, and Pakistan (admittedly, with imperfect studies) already show that wages are more strongly associated with years of schooling and labour market experience rather than test scores.

Should I stop kidding myself, or is there a way forward?

One possible takeaway from these results is that the systems are stacked against the poor. We can work at the margins but

few of the things we are doing are going to fundamentally change their life circumstances.

But perhaps that is too pessimistic. Societies can, and do, change.

Consider, for example, the experience of the U.S., which in 1933 looked very much like low-income countries today with a high weight on family background and little weight on academic performance (Hendricks, Herrington and Schoellman 2017). Yet, by 1960 family had receded to the background with academic performance gaining importance. By 1970, this trend strengthened further and for later cohorts, academic performance became the sole determinant of college attendance. The authors argue that much of this had to do with a surge in demand for college and the decreasing cost of college (in part due to subsidies for WW2 veterans through the G.I. Bill), which increased competition for high quality colleges, and allowed them to shift strategies from recruitment to selection.

Or, look more closely at India where Sam Asher has painstakingly computed educational mobility rates for India from 1950 to 1990. He shows that, even as overall educational mobility in India remained the same over this period, the mobility of scheduled castes (and to some extent, scheduled tribes) increased quite dramatically. One policy that could explain this improvement is reservations for these groups, both in higher education and in jobs.

Both examples show that in a (relatively) short time, societies can create structures that value academic achievement over family background.

But both examples also highlight that these require ongoing, systemic changes.

Ongoing because in the U.S., even as academic achievement became critical for college attendance, the overall importance of family background declined only slightly. Learning gained centre stage, but the correlation of family background and academic performance strengthened sufficiently to ensure that the children of the advantaged remained advantaged themselves. The battleground shifted, but the battle was not won.

Systemic because expanding these policies and implementing reservation policies across India has been a huge political battle with countless lives lost in the process. The rise in educational mobility for these groups has come only because scheduled caste and scheduled tribe populations mobilised and fought tooth and nail for a modicum of dignity in a country where most gains have accrued to the rich.

Ultimately, how much schooling children receive depends on how that schooling is valued in the labour market and how costly it is to obtain. In Pakistan, perhaps poor parents know that even if their very smart children continue in school, there will be no return in the labour market. Similarly, rich parents may already have a job lined up for their children – as long as they pass the critical high-stakes Grade 10 exam that opens up multiple opportunities in the public and private sectors. Or, perhaps the money is so tight that the poor cannot afford to send their children to school beyond a certain age.

At this point we don't have the answers to these critical questions. But if we continue to believe that increasing learning in schools will be sufficient to enhance equity, we are kidding ourselves.

This note is based on ongoing work with Andres Yi Chang (World Bank) and Abhijeet Singh (Stockholm School of Economics). The data on Pakistan used for this project come from the LEAPS

project and its long-term follow-up, funded by The World Bank's SRP programme and the RISE programme. I thank Lee Crawford, Shanta Devarajan, Lant Pritchett and Zainab Qureshi for detailed comments on a previous draft. The findings, interpretations, and conclusions expressed in this paper do not necessarily represent the view of the World Bank, its executive directors, or the countries they represent. I declare no relevant or material financial interests related to the material presented in this note. Correspondence to jdass1@worldbank.org.

References

- Asher, S., Novosad, P., Rafkin, C., (2018), Intergenerational Mobility in India: New Estimates from Administrative Data.
- Carneiro, P. and J. Heckman. (2003). 'Human Capital Policy', in *Inequality in America: What Role for Human Capital Policies?* ed. J. Heckman and A. Krueger (Cambridge, MA: MIT Press, 2003).
- Hendricks, L., Herrington, C., and T. Schoellman. (2017). The Changing Roles of Family Income and Academic Ability for US College Attendance. No. 1602.
- Narayan, A., Van der Weide, R., Cojocar, A., Lakner, C., Redaelli, S., Mahler, D.G., Ramasubbaiah, R., Gupta N. and S. Thewissen. (2018). 'Fair Progress?: Economic Mobility Across Generations Around the World.' *Equity and Development Series*. © World Bank Group, Washington D.C.

Familiarity does not breed contempt: generosity, discrimination and diversity in Delhi schools

– Gautam Rao

American Economic Review (forthcoming)

[Published version](#)

[Working paper version \(free\)](#)

Commentary by Susannah Hares

Want to build a better society? Make rich schools accept poor kids

Education policymakers care about more than just test scores. They probably care a lot about making policies that will help them get re-elected. They might care about particular people or places that have been historically disadvantaged. And perhaps they care about building a more integrated society: breaking down social barriers by putting children from different socioeconomic backgrounds in the same classrooms and positively influencing interracial or interclass attitudes and social behaviour.

A number of policies have been designed to do just this. For example, the landmark 1954 Supreme Court decision to desegregate American schools – decreeing that “separate educational facilities are inherently unequal” – led to court-ordered desegregation of US schools between the 1960s and

1980s. And it worked. Rucker Johnson's 2015 long-run study on mandated desegregation followed students born between 1950 and 1975 through to 2013, finding that desegregation significantly improved the educational attainment, earnings, health and incarceration rates of black students. At the same time, it had no effect on outcomes for white students. But while the legal battle for desegregation has largely been won, the reality is that schools in America today are actually becoming more segregated, mostly driven by urban housing disparities. Almost half of black and Hispanic children attend a school that is 90 per cent non-white.

The world's largest desegregation effort?

Perhaps the largest ever attempt to integrate schools is currently underway in India. Clause 12 of India's 2009 Right to Education (RTE) act mandates all private schools to reserve at least a quarter of their places for children from disadvantaged families. The government then reimburses the fee directly to the school – up to a maximum of the level of government spending per child. Schools are not permitted to select students by ability or socio-economic background. If schools are oversubscribed, they are required to run their admissions through a lottery.¹

With over 60 million children estimated to be enrolled in private schools in India – and potential consequences for more than 300 million school-age kids – this is a massive effort by the government to give disadvantaged children access to the same educational opportunities as their richer peers. It is an explicit acknowledgement of the acute

¹ While fair in theory, this does require parents to be able to fill out and submit the application form and obtain documentation proving their eligibility. Therefore children whose parents want to invest in their education, have knowledge of the programme and have the time and ability to complete the paperwork are most likely to benefit.

segregation in India's education system, with the rapid rise in private education reflecting and indeed amplifying societal and economic inequalities. RTE Clause 12 – in theory at least – is a direct challenge to some of the institutions and vested interests that propagate the segregation of educational opportunities available to Indian children.

Bringing poor kids into elite classrooms makes the richer kids nicer ...

RTE Clause 12 is by no means an unmitigated success. Economists Esther Duflo and Abhijit Banerjee have criticised the policy (2011) for its focus on inputs and process over learning outcomes. A number of studies, including by Geeta Gandhi Kingdon and Francis Teal (2010) and Abhijeet Singh (2013), show that the criteria used for official registration of private schools – physical infrastructure, teacher qualifications etc. – are not good predictors of school performance. But a recently published paper by Gautam Rao (2013) sheds light on an isolated but important effect of the policy: mandating elite schools to admit poor kids into classrooms full of rich children has inherent value in terms of social cohesion and reducing prejudice.

Rao's study exploits the phasing-in of the policy in 17 "elite" private schools in Delhi (schools with fees of more than 2,000 rupees (\$30) per month). To be eligible for the government subsidy, children had to come from households with incomes of less than 100,000 rupees per year (\$1,500) – around the 45th percentile of household income distribution.² The average "poor" student in Rao's sample is from the 25th percentile. In contrast, the typical "rich" student in the sample

² This supports concerns that the RTE private school subsidy policy has not achieved its objectives of providing opportunities to children from the poorest quintile.

is from well above the 95th percentile of income distribution. So, while it's possible to dispute how pro-poor the RTE policy really is, the poor kids in this experiment are – without question – of a very different socioeconomic status to the wealthy kids enrolled in the study's elite private schools.

Rao gathered attainment and administrative data on around 2,700 kids and compared those in year groups exposed to Clause 12 to those in immediately older cohorts who were not. The results are nothing short of heart-warming. Being around their poorer classmates made wealthier students more pro-social. Rich kids in the treatment group were 60 per cent more likely to volunteer for charity. It made them substantially more generous – they were 45 per cent more willing to share money with poorer kids. They wanted to socialise with poor children – they were much more enthusiastic to accept weekend playdates kids from schools in poorer neighbourhoods. And, it made the rich kids more willing to pick poor children to be on their side in team sports. This is not only a pro-social behaviour, but one that corrects for a potentially *inefficient* behaviour: i.e. picking low-ability rich kids over high-ability poor kids.³ Integration in schools reduces this inefficiency – high-ability poor children become more likely to be chosen. Notably, these effects were not just limited to poor kids at their school. Mixing classrooms changed the behaviour of rich kids toward other poor children – strangers to them – outside their school.

... and doesn't have negative effects on their attainment or behaviour (although they do swear a bit more)

A common pushback to desegregation is the claim that bringing in disadvantaged children will have negative effects on children already in elite schools. There is a substantial but

³ Measured through a "dictator games" in a lab setting.

inconsistent literature examining peer effects in education to draw from. For example, Hanushek et al (2003) found that peer achievement has a significant positive effect on achievement growth, with students throughout the score distribution benefiting from higher achieving classmates, while Angrist and Lang (2004) found only modest and short-lived effects of desegregation on test scores.

Returning to Delhi, Rao found that – while poor students did perform worse than rich students in tests – bringing them into the classroom had no significant effect on their rich classmates' overall scores. There was a marginal negative effect on English language scores,⁴ which also happens to be the subject with the greatest achievement gap between rich and poor. This most likely reflects the fact that the poor kids in the sample almost exclusively speak Hindi at home.

Rao found no increase in disruptive or violent behaviour in the classrooms he studies. However, it does seem that the incoming kids from disadvantaged families taught their affluent classmates some new language skills. The number of rich kids using swear words increased by seven percentage points in the treatment group.

Education policy-makers need to take desegregation more seriously

Rao's findings are somewhat reflected at tertiary level in Colombia in a study by Juliana Londono-Velez. She studies a government scholarship programme that allowed 10,000 low-income students to attend an elite university. Like Rao, she found that mandated classroom integration made richer kids

⁴ The non-effect on maths and Hindi scores is consistent with findings from Muralidharan and Sundararaman (2013) who found no effect on the achievement of existing students in private schools in rural India when poorer and lower-achieving children entered their classroom on vouchers.

nicer. Exposure to their poorer classmates made them more sympathetic to challenges of poverty and social mobility, increased their support for meritocracy in college admissions, and increased support for wealth redistribution. The policy had no negative effect on academic outcomes: grades and drop-out rates were unaffected.

These studies are of great significance for both poor and rich countries, particularly at a time when policy-makers are grappling with integration of refugees into school systems, financing scale-up of post-primary education, and a trend of **increasing not decreasing segregation** of schools. It may seem like common sense to assume that getting rich and poor kids to go to the same schools is a good way to equalise educational opportunities and improve social cohesion. Yet, government-mandated integration of schools – through policy instruments like India’s RTE Clause 12 – is rarely discussed in education policy-making. The policy goals in the World Bank’s **framework paper** on “what matters most for equity and inclusion in education systems” make no reference to school desegregation and integration. And vested interests in **developed** and developing countries alike seem intent on protecting and even increasing highly selective admissions into elite government schools.

Rao’s study makes a substantial contribution to the literature on desegregation and peer effects in education. It should be on the reading list of every education policy-maker. He shows that government-mandated desegregation policies can have a profound impact – beyond learning outcomes – on social cohesion, social behaviour, and perhaps even nation-building. For all its flaws, RTE Clause 12 has the potential to help address the socioeconomic stratification in Indian schools that has perpetuated inequality over generations. Of course, test scores

really matter. But in India and elsewhere – perhaps now more than ever – so do nicer schools and more inclusive societies.

With thanks to Gautam Rao and Justin Sandefur for their comments.

References

- Angrist, J.D. and K. Lang, K. (2004). ‘Does School Integration Generate Peer Effects? Evidence from Boston’s Metco Program.’ *American Economic Review*, 94(5): 1613-1634.
- Banerjee, A., and E. Duflo. (2011). *Poor economics: A radical rethinking of the way to fight global poverty*. New York: Public Affairs.
- Hanushek, E.A., Kain, J.F., Markman, J.M., and S.J. Rivkin. (2003). ‘Does peer ability affect student achievement?’ *Journal of Applied Economics* 18, 527-544.
- Johnson, R.C. (2011). ‘Long-run impacts of school desegregation & school quality on adult attainments’ (No. w16664). National Bureau of Economic Research.
- Kingdon, G., and F. Teal. (2010). ‘Teacher unions, teacher pay and student performance in India: A pupil fixed effects approach.’ *Journal of Development Economics*, 91(2), 278-288.
- Muralidharan, K. and V. Sundararaman. (2015). ‘The aggregate effect of school choice: Evidence from a two-stage experiment in India.’ *The Quarterly Journal of Economics*, 130(3), 1011-1066.
- Rao, G. (2013). ‘Familiarity does not breed contempt: Diversity, discrimination and generosity in Delhi schools.’ *Job Market Paper*.
- Singh, A. (2013). ‘Size and sources of the private school premium in test scores in India.’ *Young Lives*.



**Cambridge
Assessment**



CENTRE FOR EDUCATION ECONOMICS

23 Great Smith Street, Westminster, London SW1P 3DJ

www.cfee.org.uk
comms@cfee.org.uk

£8.99

