Public Disclosure Authorized

Enhancing the Quality of Public Service Delivery:
Insights from Recent Research

Michael Woolcock

Good policies are necessary but insufficient for delivering high quality public services. Building organizations capable of implementing increasingly complex policies at scale is crucial. But most developing countries lack such organizations and – alarmingly – most are not improving. Effective reforms in China, Singapore, and Vietnam, however, and from unlikely service delivery successes elsewhere, suggest that three key factors are essential to building implementation capability: beginning with problems nominated and prioritized by local staff; promoting iterative experimentation to identify a range of plausible alternatives; and sharing emerging solutions through a frontline community of practice. These factors also suggest there is much to be gained from embedding researchers into delivery systems themselves, the better to understand context, to map and explain variation, and to facilitate real-time decision making in response to changing realities.

Most development research papers and debates seek to discern the implications of empirical findings for "policy," the underlying assumption being that a change in policy will lead to desirable outcomes. But even if a specific shift in policy is clearly articulated, technically sound, and broadly supported, it cannot be assumed that the government can implement it, especially in low- and middle-income countries. For example, most governments in such countries have policies to minimize (if not eliminate) corruption, but too often such practices continue unabated. Indeed, one could say that a defining characteristic of developing countries is their low capability to implement complex policies, and that policy failure may often stem less from the policy's design than its weak implementation.

Beyond policy, issues of implementation can be an important focus of research: How well do governments around the world implement their policies? Why are some so much better at it than others? Why are some improving more rapidly than others? Why are some policies – or certain aspects of a policy – more difficult to implement than others? What can be done to improve capability for implementing public policy?

Answers to these questions are beginning to become clearer, especially in the field of public service delivery, where implementation quality clearly matters: we all benefit from clean streets and good schools, we all fear having an illness be misdiagnosed. Overseeing such services (if not necessarily delivering them) is a government's core responsibility, and the realm in which citizens and states most frequently encounter one another. How well a government implements key services determines whether lives are saved or lost, whether citizens have a hopeful or bleak future, and whether elections are won or lost. There are powerful moral and instrumental reasons for "getting implementation right" in all countries, but especially in developing ones.

What insights does recent research yield?

First, the unhappy news. As table 1 shows, the overall state of implementation capability around the world – as best one can measure such vexing issues – is mostly low in most developing countries and mostly declining. Recent analyses suggest that only one major country outside the Organisation for Economic Co-operation and Development (OECD) – Indonesia – is on track to become a "high capability" state – a state that attains the level of the lowest OECD country – by the end of the twenty-first century. At current rates, many of today's low- and middle-income countries will never reach high capability status (see Andrews, Pritchett, and Woolcock 2017).

The picture is little better viewed from the perspective of national data. In recent years, for instance, successive cohorts of Indian primary school students have performed worse on tests of learning (Pritchett 2014). Even returning misaddressed mail – perhaps the least difficult task that 157 governments around the world have formally committed themselves to doing – seems beyond many states, including long-established ones (such as Russia and the Arab Republic of Egypt) (Chong et al. 2014). Anthropologists studying the roll-out of pioneering social programs to support poor elderly women lament that noble policy intentions often fail to be realized because, in the poorest areas, seemingly unambiguous personal facts – age, land ownership (a proxy for wealth), number of children – are unable to be administratively verified, leaving final program eligibility decisions in the hands of frontline administrators, too many of whom abuse this discretion, in the process becoming, for these poor elderly women, part of the problem rather than part of the solution (Gupta 2012).

But the news is not all bad; indeed, there is much to celebrate and much to learn from in recent studies of effective public service delivery implementation. A defining characteristic of complex policy challenges is that the

WORLD BANK GROUP
Global Knowledge & Research Hub in Malaysia

Table 1. The (Unhappy) State of State Capability around the World

	Rapid negative	Slow negative	Slow positive	Rapid positive
Rate of change	g < -0.05	-0.05 < g < 0	0 < g < 0.05	g > 0.05
Initial levels of State Capability (SC)				
Strong SC > 6.5		BHR, BHS, BRN	**CHL (0), SGP (0), KOR (0), QAT (0)**	ARE (0)
Total: 8	0	3	4	1
Medium 4 < SC < 6.5	MDA, GUY, IRN, PHL, LKA, MNG, ZAF, MAR, THA, NAM, TTO, ARG, CRI	PED, EGY, CHN, MEX, LBN, VNM, BRA, IND, JAM, SUR, PAN, CUB, TUN, JOR, OMN, MYS, KWT, ISR	KAZ (10,820), GHA (4,632), UKR (1,216), ARM (1,062), RUS (231), BWA (102), **IDN (68), COL (56), TUR (55), DZA (55), ALB (42), SAU (28), URY (10), HRV (1)**	
Total: 45	13	18	14	0
Weak 2.5 < SC < 4	GIN, VEN, MDG, LBY, PNG, KEN, NIC, GTM, SYR, DOM, PRY, SEN, GMB, BLR	MLI, CMR, MOZ, BFA, HND, ECU, BOL, PAK, MWI, GAB, AZE, SLV	UGA (6,001), AGO (2,738), TZA (371), BGD (244), ETH(103), ZMB (106)	
Total: 32	14	12	6	0
Very weak SC < 2.5	YEM, ZWE, CIV	SOM, HTI, PRK, NGA, COG, TGO, MMR	SDN (7,270), SLE (333), ZAR (230), IRQ (92)	NER (66), GNB (61), LBR (33)
Total: 17	3	7	4	3
TOTAL: 102	**30**	**40**	**28**	**4**

Source: Adapted from Andrews et al, 2017, 20. Calculations come from rescaled indicators of state capability derived from three sources (from 2015): *Quality of Government* (http://qog.pol.gu.se/); *(Failed) Fragile States Index* (http://fsi.fundforpeace.org/); and Worldwide *Governance Indicators* (http://info.worldbank.org/governance/wgi/index.aspx#home).
Note: The number in brackets after some countries is the years it will take for that country to have a high capability state (a state at the level of the lowest OECD country). Countries in bold are "historically developing countries." Only 13 of them are on a path to having a "high capability" state by the end of the twenty-first century. In those countries without a number, state capability is declining. g = rate of change; SC = state capability.

quality of their realization almost always varies greatly: such policies and programs — even when carefully designed and adequately supported — are likely to work well in some places and for some people, to have little effect on others, while for some they may even make things worse. Thus, the empirical challenge is to map and explain this variation, and to use insights gleaned from the analysis to guide possible strategies for improvement elsewhere. A recent study of education and health care across the Middle East and North Africa, for example, showed that staff absenteeism at health clinics in the Republic of Yemen ranged from 8 percent to 92 percent. The policy in Yemen, of course, is that staff should show up for work each day, yet there is clearly wide variation in how well that policy is realized. In the highly challenging context of West Bank and Gaza, some schools are nonetheless able to teach students well enough that they score at levels comparable to average schools in OECD countries. How do they do this? By cultivating a leadership culture that engenders high trust and cooperation. The principal, students, teachers, parents, community members, and mid-level administrative officials jointly nominate and prioritize their problems, and work together to identify and implement locally legitimate solutions (see Brixi, Lust, and Woolcock 2015).
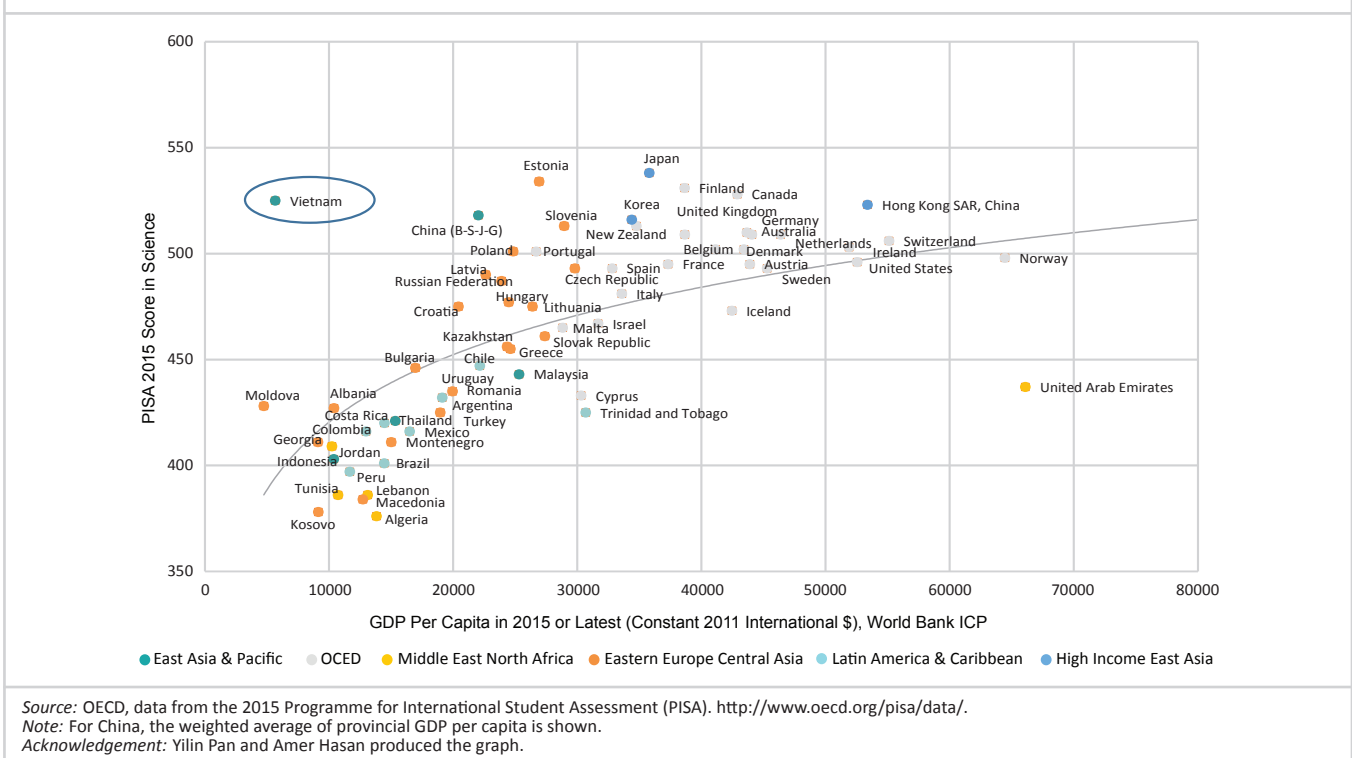
In East Asia, the effectiveness with which service delivery policies are implemented also varies widely, between and within countries. Singapore, for example, has risen to the top of most global measures of state capability despite being a very poor country less than 40 years ago. A recent study of the policy implementation strategies adopted by Singapore argues that while it identified clear aspirational goals, it nonetheless encouraged frontline officials to experiment and iterate (Goh 2016). Similarly, a study of the reform process in contemporary China during its rise to global economic powerhouse shows that local officials, far from being required to forcefully apply a rigid singular strategy, were instead given wide autonomy to use their personal networks and professional skills to promote investments, expand economic opportunities, and identify local solutions to local problems (Ang 2016).

Perhaps the most spectacular success in service delivery attainment has been Vietnam's education performance. In standardized international examinations conducted in 2015, students in Vietnam attained scores higher than those in Switzerland, although Vietnam has a per capita income level roughly one-tenth the size (see figure 1). Researchers are still seeking to explain exactly how this has been attained, but for the moment it seems there is agreement that the familiar "structural determinants" of such outcomes elsewhere – levels of spending, class sizes, teacher education, demographic characteristics – cannot account for it (Dang 2017). However, even if it is likely that Vietnam's current success can be attributed to its own version of the discretionary strategy used in Singapore and China, caution should be exercised in inferring that particular programmatic details from these countries warrant the status of global "best practices" that can be readily adopted elsewhere, with expectations of similarly impressive results (Bold et al 2013). Such "cut and paste" reform strategies to improve the quality of service delivery – in which the visible characteristics of successful approaches (legislation, organizational charts, training manuals) rather than the context-specific processes and practices that generated them – have a long but inglorious history (Andrews 2013). In many instances, adopting such "best practice" reforms have enabled leaders to superficially imbue their actions with legitimacy and thereby claim "success," despite the fact that so often they have achieved nothing of the sort – precisely because actual capability for policy implementation is not (and has never been) acquired this way (Andrews, Pritchett, and Woolcock 2017).

Expanding roles for research in enhancing service delivery quality

The process by which research on service delivery effectiveness is being conducted is undergoing innovations at multiple levels, from field experiments (see Banerjee, Duflo, and Glennerster 2008), to systematic reviews (Snilstveit et al. 2016) and "big data" analyses, to case studies (exemplified by the Global Delivery Initiative and Princeton University's Innovations for Successful Societies), household surveys, and intensive qualitative investigations (conducted by research teams embedded within development projects). The complex nature of service delivery challenges means that, ideally, a diverse but integrated array of methods and data is needed to enhance the quality of decision making. But a defining characteristic of effective service delivery implementers is that they proceed through iteration. Frequent feedback from frontline staff and citizens (who in turn are given discretion, space, and resources to explore a range of plausible options) can be used to make real-time adjustments to design and practice. This process would suggest giving priority to research methods optimally placed to facilitate – and ensure the quality of – such strategies based on iteration. It would also suggest that, just as there has been a revolution in recent years in the sophistication of tools to evaluate projects, there needs to be a corresponding refinement in the tools used to monitor projects, shifting their role from that of compliance to learning. In short, broader approaches to both 'M' (monitoring) and 'E' (evaluation) are needed if researchers are to help implementers achieve their goals (see Pritchett, Samji, and Hammer 2013).

Figure 1. Some Countries (like Vietnam) Provide Education Much Better than Others

Source: OECD, data from the 2015 Programme for International Student Assessment (PISA). http://www.oecd.org/pisa/data/.
Note: For China, the weighted average of provincial GDP per capita is shown.
Acknowledgement: Yilin Pan and Amer Hasan produced the graph.

Rao's work with the Social Observatory in India exemplifies what such an approach might look like in practice. Drawing on approaches from across the methodological spectrum, research teams work closely with project implementers to discern not just "whether" a given intervention is working on average, but "how" and "for whom" (see Ananthpur, Malik, and Rao 2014; Joshi and Rao 2017). It is often not at all obvious what changes need to be made over time and across contexts to improve project effectiveness, and most frontline staff do not have the luxury of waiting three years for empirical findings from "gold standard" evaluations to make decisions. Thus, deploying a mixed methods research approach to generate "good enough" evidence in real time is likely to be the most useful (Bamberger, Rao, and Woolcock 2010). Similar interdisciplinary initiatives to enhance the quality of service delivery are being pioneered in Indonesia (Alatas 2017). Dedicated "delivery units," such as those being pioneered in middle-income countries such as Malaysia, can also play an important role in rapidly sharing specific insights on innovative problem-solving strategies across professional communities of practice.

Conclusions, implications, and applications

Development challenges in the coming decades will require governments to implement increasingly complex and (potentially) contentious policy agendas, at scale. Administrative systems that are already struggling will have to accommodate populations that continue to grow in size and diversity, and whose expectations of their governments – especially in education – will only rise as the demands of an increasingly connected and competitive world becomes apparent (Kim 2017). One reading of the prevailing evidence would lead to rather pessimistic conclusions regarding the likelihood that these expectations can be met. But a different reading would suggest that researchers, practitioners, and policymakers have an opportunity to respond to these challenges in creative ways, forging innovative strategies built on new understandings of how organizations actually acquire effective implementation capability.

These strategies have an important place for research, though one centered less on identifying universal "best practices" (as determined by international experts) that can be replicated and scaled, and focused more on understanding the forms and sources of variation from context to context. Such insights can help identify how and where local solutions have emerged in response to local problems, and promote the sharing of these insights with others. Doing this well requires trusting mid-level and frontline staff to find and amplify the successes that someone somewhere somehow has already achieved.

Note. This Research & Policy Brief summarizes insights drawn from a research conference on "Enhancing the Quality of Service Delivery," hosted in Kuala Lumpur in January 2017 by the World Bank's Knowledge and Research Hub in Malaysia.

References

Alatas, Vivi. 2017. "Service Delivery to Reduce Inequality." Paper presented at the "Enhancing the Quality of Service Delivery" conference, World Bank Knowledge and Research Hub, Kuala Lumpur, Malaysia, January 16.

Ananthpur, Kripa, Kabir Malik, and Vijayendra Rao. 2014. "The Anatomy of Failure: An Ethnography of a Randomized Trial to Deepen Democracy in Rural India." Policy Research Working Paper 6958, World Bank, Washington, DC.

Andrews, *Matt. 2013. The Limits of Institutional Reform in Development: Changing Rules for Realistic Solutions.* New York: Cambridge University Press.

Andrews, Matt, Lant Pritchett, and Michael Woolcock. 2017. *Building State Capability: Evidence, Analysis, Action.* New York: Oxford University Press.

Ang, Yuen Yuen. 2016. *How China Broke the Poverty Trap.* Ithaca, NY: Cornell University Press.

Bamberger, Michael, Vijayendra Rao, and Michael Woolcock. 2010. "Using Mixed Methods in Monitoring and Evaluation: Experiences from International Development." In *Handbook of Mixed Methods in Social and Behavioral Research,* 2nd revised edition, edited by Abbas Tashakkori and Charles Teddlie, 613–41. Thousand Oaks, CA: Sage Publication.

Banerjee, Abhijit V., Esther Duflo, and Rachel Glennerster. 2008. "Putting a Band-Aid on a Corpse: Incentives for Nurses in the Indian Public Health Care System." *Journal of the European Economic Association* 6 (2–3): 487–500.

Bold, Tessa, Mwangi Kimenyi, Germano Mwabu, Alice Ng'ang'a, and Justin Sandefur. 2013. "Scaling Up What Works: Experimental Evidence on External Validity in Kenyan Education." Working Paper 321, Center for Global Development, Washington, DC.

Brixi, Hana, Ellen Lust, and Michael Woolcock. 2015. *Trust, Voice and Incentives: Learning from Local Success Stories in Service Delivery in the Middle East and North Africa.* Washington, DC: World Bank.

Chong, Alberto, Rafael La Porta, Florencio Lopez-de-Silanes, and Andrei Shleifer. 2014. "Letter Grading Government Efficiency." *Journal of the European Economic Association* 12 (2): 277–99.

Dang, Hai-Anh. 2017. "High Quality Education Despite Resource Constraints? The Case of Vietnam." Presentation at World Bank Policy Research Conference on "Enhancing the Quality of Service Delivery," Kuala Lumpur, Malaysia, January 15-16.

Goh, Nigel. 2016. *Detours, Dead-Ends and Diversions: Singapore's Road to Development Reconsidered.* Singapore: UNDP Global Centre for Public Service Excellence.

Gupta, Akhil. 2012. *Red Tape: Bureaucracy, Structural Violence and Poverty in India.* Durham, NC: Duke University Press.

Joshi, Shareen, and Vijayendra Rao. 2017. "Who Should Be at the Top of Bottom-Up Development? A Case Study of the National Rural Livelihoods Mission in Rajasthan, India." Policy Research Working Paper 7996, World Bank, Washington, DC.

Kim, Young Eun. 2017. "The Role of Education in Enhancing Productivity: Some International Comparisons." Presentation at World Bank Policy Research Conference on "Enhancing the Quality of Service Delivery," Kuala Lumpur, Malaysia, January 15–16

Pritchett, Lant. 2014. "The Risks to Education from Design Mismatch and Global Isomorphism." WIDER Working Paper 2014/039, United Nations University.

Pritchett, Lant, Salimah Samji, and Jeffrey S. Hammer. 2013. "It's All about MeE: Using Structured Experiential Learning ('e') to Crawl the Design Space." Faculty Research Working Paper 13-012, Harvard Kennedy School, Cambridge, MA.

Snilstveit, Birte, Jennifer Stevenson, Radhika Menon, Daniel Phillips, Emma Gallagher, Maisie Geleen, Hannah Jobse, Tanja Schmidt, and Emmanuel Jimenez. 2016. "The Impact of Education Programmes on Learning and School Participation in Low- and Middle-income Countries' Systematic Review." Summary 7, International Initiative for Impact Evaluation, New Delhi.

WORLD BANK GROUP
Global Knowledge & Research Hub
in Malaysia