Comment: Scalability cannot be the sole criterion for policy decision making

Moses Oketch

Crawfurd, Hares, and Sandefur raise several important issues related to what works at scale in terms of education policy and interventions. My comments here focus on two key issues, of relevance to low- and lower-middle-income countries: (1) the need for renewed attention to the unfinished business of access and (2) the complexity of scaling pedagogic reforms shown to work in the pilot phase.

First, concerning the need for renewed attention to the unfinished business of access. The focus on learning instead of schooling is central to realizing Sustainable Development Goal 4. Without learning, there is little point in going to school; without children going to school, it is nearly impossible to organize formal learning at scale. Parents expect schooling to lead to learning. However, the "learning crisis" changes the terms of the debate since it is largely presented and interpreted as an argument and movement that is against continued expansion in enrollment and attainment. As Pritchett and Sandefur (2020) have noted, it will require both universal schooling and a dramatic improvement in learning profiles to achieve the SDG targets. However, dramatic improvement in learning is not going to happen easily and analysis comparing countries' learning profiles does not show decisively what works, but the clear message in the learning crisis movement is that schooling itself without evidence that it also generates sufficient learning is not good enough. In order to dramatically improve learning, Crouch, Kaffenberger, and Savage (2021) have argued that there is a need to focus on systems improvement, and to use foundational learning as the guiding principle to ratchet up learning. However, Crawfurd, Hares, and Sandefur challenge that view, at least in low- and lower-middle-income countries with space to expand access. They note that pedagogy reforms can be difficult to implement and there is not enough evidence of their scalability in government school systems. Expecting imperfect government systems to dramatically improve learning has proven to be a tall order. So, they argue that expanding access, providing school meals, and extending school time are policies that have worked at scale, and focusing on these is an actionable agenda that can raise learning outcomes in the near time. Their main message is that education pays even where schooling does not generate stellar learning outcomes; therefore, from an economic standpoint, even rudimentary schooling may be a very good investment. The main theme of their argument is that scalability of a program has to be critical when interventions are presented to a minister of finance. While this is a reasonable argument, it is also very narrow. I would argue that scalability cannot be the sole criterion for policy decision making, although it may be one of the considerations.

Second, concerning the complexity of pedagogic reforms and scalability, there is general agreement that learning needs to improve dramatically in low- and lower-middle-income countries, and even Crawfurd, Hares, and Sandefur do not argue against this. Most commonly for low- and lower-middle-income countries, there is agreement that something must be done about the learning crisis. The success in schooling should not be allowed to go to waste. Thus, I would argue that the choice of Type A policies—those that are more effective when well implemented but require highly skilled staff (these policies include structured pedagogy, teacher coaching, or home visits for early child development)—versus Type B policies—those that might be less effective when implemented in terms of improving learning but are more robust to weak implementation (these include school-building, lengthening the school day, or providing school meals)—is a false choice. Instead, both Type A and Type B policies are needed in a country, as addressing the learning crisis requires education system improvement. I have argued elsewhere (Oketch 2019), as Crawfurd, Hares, and Sandefur do in their chapter, that measures of performance, efficiency, and effectiveness often embedded and dominant in randomized control trial (RCT) studies do not provide explanations of how and why an education system "is where it is" or of "what works" to improve it. But, I would also argue that RCT-based studies may offer some insight at small scale on potential mechanisms of change, which can help to identify where the "blockage" lies at the macro level, even when these pilots have not proven scalable. So, while scalability may be a useful criterion for policies and indeed critical, there are still many systems-improving lessons that can be learnt from pilots and projects that haven't proven scalable. What I would argue against, and where I agree with Crawfurd, Hares, and Sandefur's argument, is that RCTs and pilot projects that have demonstrated success should not crowd out those programs that have already shown success when rolled out at scale, and a gradualist approach rather than a big bang rollout of pedagogy reforms might be a wise approach to present to a minister of finance. It is not obvious that systems that have improved learning have done so through relying on pilots, but they have certainly learnt gradually or in an evolutionary manner.

References

Crouch, L., M. Kaffenberger, and L. Savage. 2021. "Using Learning Profiles to Inform Education Priorities: An Editors' Overview of the Special Issue." *International Journal of Educational Development* 86, 102477.

Oketch, M. 2019. "Randomized Controlled Trials: Limitations for Explaining and Improving Learning Outcomes." In *World Yearbook of Education 2019 Comparative Methodology in the Era of Big Data and Global Networks*, edited by R. Gorur, S. Sellar, and G. Steiner-Khamsi. London: Routledge.

Pritchett, L., and J. Sandefur. 2020. "Girls' Schooling and Women's Literacy: Schooling Targets Alone Won't Reach Learning Goals." *International Journal of Educational Development* 78, 102242.