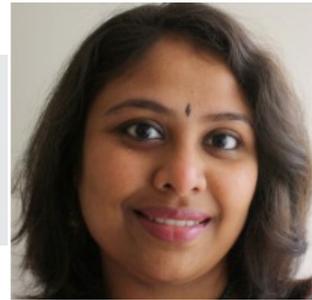


Details Matter: An Evaluation of ‘Evaluations’

Harini Kannan



Introduction

In August 2009, the Indian parliament enacted the Right to Education (RTE) Act which enshrined education from 6 to 14 years as a ‘right’. The Act additionally mandated a variety of ‘requirements’ relating to infrastructure, Pupil Teacher Ratio (PTR), curriculum, teacher training, inclusionary education, and the focus of this article – a continuous and comprehensive student evaluation system (CCE). The objective of these mandates was ensuring a ‘quality’ education for children.

Amongst the mandates, the introduction of CCE signaled a paradigm shift in India’s public education. Historically, student evaluations have focused on measuring academic knowledge gained by children over the course of a term or year through terminal examinations. These were high-stakes exams, as scores carried a lot of weight in determining whether students were promoted or detained in the same standard. Traditional systems of evaluation were also very narrow in their scope – the focus was primarily on evaluating students on ‘subject’ knowledge and very rarely, if at all on other aspects. CCE on the other hand, was meant to be both continuous and comprehensive: by ‘continuous’ it meant that evaluation of students would be done ‘continually’ over the course of the academic term, and by ‘comprehensive’ it meant that evaluations would not just focus on learning in academic subjects, but also on co-curricular activities and behaviour. The underlying sentiment impelling these changes was that schooling should foster learning, be enjoyable and less stress inducing and focus on the holistic development of a child.

As the CCE programme was being rolled out, researchers affiliated with the Abdul Latif Jameel Poverty Action Lab (J-PAL) conducted a rigorous evaluation of the implementation of the CCE programme in Haryana.¹ The results were

sobering. In spite of the promise of CCE to catalyse improvement in primary schools, the programme does not appear to be meeting the basic objective of increasing learning outcomes.

‘Theory’ of CCE

The ‘continuous’ aspect of CCE is well grounded in student evaluation theory. Continuous evaluation typically consists of ‘formative’ and ‘summative’ evaluations which are carried out throughout the term and at the end of the term respectively. Formative evaluations are diagnostic in nature and the information gathered from them is used to strengthen the teaching-learning process. Summative assessments, conducted at the end of the term, enable quantification of students’ gains in knowledge. While summative assessments are typically pen and paper tests, formative assessments can range from informal ones such as oral questions asked during class to pop quizzes and projects. These types of evaluations are therefore designed to be complementary to one another and more inclusive to students who may face challenges with respect to traditional assessments.

The ‘comprehensive’ aspect of CCE focuses on holistic development of children. This has been emphasised by educators over time as essential to ensure that students gain key life-skills and be productive members of society. However, despite this lofty sentiment, there has been no formal integration of measures to cultivate these life-skills in the standard school curriculum; rather this has been left to schools and teachers. Comprehensive evaluations have been seen as a tool to align the focus of the goal with the processes to achieve the same.

Implementation and Evaluation of CCE

While the RTE Act mandated the introduction of CCE, it stipulated that states design their own programmes as suitable for their local needs. The

¹This paper is a non-technical summary of Berry et. al. (2018). Please review the paper for more details on the evaluations and results. The working paper can be accessed at https://www.povertyactionlab.org/sites/default/files/publications/Failure-of-Frequent-Assesment_Berry-et-al._feb_2018.pdf

Central Board of Secondary Education (CBSE) was the pioneer in introducing CCE in affiliated schools, and its programme has served as the blueprint for various states developing their own programmes. As of April 2018, almost all states have implemented CCE, though there are variations in the coverage across schools and standards.

In spite of CCE's widespread adoption as well as support from educationalists (CBSE 2009), little is known about the program's effects on outcomes it is supposed to help achieve. Quantification of a programme's impact is ultimately an empirical question and one cannot conclude on the 'impact' of a program without subjecting it to a rigorous impact evaluation. Impact evaluations of social programmes and policies are common in the development economics literature and increasingly adopted by governments and civil society organisations to understand whether new programmes 'work.'

Shortly after CCE's adoption by the Central Government, the Government of Haryana Department of Education reached out to J-PAL to conduct an evaluation of the CCE programme in the State of Haryana during the 2012-2013 school year. While the evaluation was limited to Haryana's CCE implementation, broader lessons can be distilled. Before the article outlines the programme, findings and conclusions; a short refresher on impact evaluations is warranted.

Impact Evaluation using Random Assignment

While 'impact' has many colloquial meanings, in the context of empirical research it has a very precise definition. Impact is defined *as the difference in outcomes (can be learning, health, economic etc.) of a group when they have been exposed to a programme compared to outcomes had they not been exposed to the program (counterfactual)*. As one immediately recognises, it is impossible to observe the counterfactual and we need to use other means to create a group to compare outcomes. In academic parlance, this group is called the comparison or control group. While there are many different methodologies that can be used to quantify impact, the credibility of the method rests primarily on how the comparison group is created. These groups can be created in a random, quasi-random or non-random manner. The J-PAL affiliated researchers chose to evaluate the impact of the CCE programme using a Randomised Controlled Trial (RCT) design. RCTs are considered the most

rigorous and credible way to evaluate programmes, as they involve randomly assigning individuals or groups to receive a program, while others are randomly assigned not to participate in the programme. The strength of the RCT design rests on random assignment of individuals or groups so that prior to the implementation of the programme, these groups are similar in nature and they only differ in their exposure to the program. Given this, if there are differences in outcomes at the end of the program, they can solely be attributed to the programme and not to other factors.

Haryana's CCE programme design and roll-out

In 2011, Haryana was one of the first states to develop and pilot their CCE program. The programme was designed by the State Council for Education Research and Training (SCERT) and was influenced heavily by CBSE's program. Continuous evaluation was operationalised by evaluating students in standards 1 to 5 on a monthly basis and standards 6 to 8 on a quarterly basis. To facilitate diagnostic evaluation, languages were evaluated on the basis of listening, reading and writing skills, while mathematics and environmental sciences (EVS) were evaluated on the basis of learning of fundamental concepts. Key sub-skills/concepts were identified for many of these, and assessment of children was required across all. The program required significant documentation in form of monthly evaluation sheets which the teacher used to record evaluation for each sub-skill including broader descriptive comments, and term-wise report cards which were used to provide a consolidated status of learning for each child enrolled in the class. Marks were eschewed and grades were provided at the end of the term to standards 6, 7 and 8, while a summary of descriptive remarks were provided for students in standards 1 to 5. The programme necessitated the use of a variety of evaluation tools such as oral recitation/Q&A, class participation, quizzes, unit tests and projects. To facilitate objectivity and standardisation in evaluation, detailed grading rubrics were provided. In addition to scholastic aspects, co-scholastic aspects such as participation, creativity and skill in cultural activities, and personal qualities such as respect, cleanliness, leadership quality, were also assessed.

Though the programme was conceptualised by the SCERT, the training for the programme was outsourced to private agencies. A 'cascade' training

model was adopted where the SCERT faculty oriented resource persons from agencies, who then trained master trainers who in turn trained the trainers. Teacher training was conducted for seven days at block headquarters by the trainers. Teachers were primarily trained on how to conduct student evaluations and complete the required documentation with some focus on how to change teaching practices or otherwise aid low performers.

During the course of designing the programme and planning for the evaluation; J-PAL researchers emphasised the need for a strong mentoring and monitoring mechanism in the field. Programme take-up and implementation often falters when there is no on-going support for implementer or participants. To ensure the CCE program was not consigned to that fate, the Education Department requested that the J-PAL research team help them set-up systems for monitoring and mentoring. Interestingly, while there was a pre-existing cadre of government officials called ABRC (Assistant Block Resource Coordinators) whose main role on paper was to support school functioning; their role as academic advisors was de-emphasised. They were instead used as 'couriers' to communicate with teachers, gather data and organise events. In consultation with the department, the J-PAL research team worked to operationalise and systematise the role of ABRC as mentors and monitors. This was done through clearly defining responsibilities, training them on the CCE programme and how to mentor and monitor it and finally setting up an internal implementation review and feedback mechanism within the district.

Evaluation Design, Sample and Data²

The Education department requested us to situate the evaluation across four blocks in two districts – Kurukshetra and Mahendragarh. Five hundred schools were sampled from a universe of all schools in the four blocks (Ateli and Narnaul in Mahendragarh and Thanesar and Pehowa in Kurukshetra). Four hundred of these schools were primary schools with standards 1 to 5, while the remaining hundred were upper primary schools consisting of standards 6 to 8 i.e. middle, high or senior secondary schools. To operationalise the RCT, the schools were randomly assigned to either receive the CCE intervention or to a control group.

Impact on what'?

While carrying on an impact evaluation, it is necessary to identify the outcomes that the program is designed to affect. Once these broad outcomes are conceptually defined, it is necessary to break them down in measurable indicators. There are a variety of outcomes that can be affected by the introduction of CCE – students may experience less stress, they may find school more enjoyable, or they may have improved self-esteem and learning outcomes. The focus of our evaluation was limited to quantifying CCE's impact on learning outcomes for one major reason. In a country like India, which has made significant strides in student enrolment, learning outcomes have not kept pace. Year after year, the National Assessment Surveys and Annual Status of Education Reports (ASER) report little, if any increases in learning outcomes. Therefore, a far reaching program such as CCE should first and foremost lead to improvements in learning outcomes. This focus was supported by senior bureaucrats who believed that, given the investment in the design and implementation of CCE; they would like to understand if the programme addressed the key issue of learning outcomes. We therefore decided to focus on learning outcomes in Hindi and Math.

Findings and rationale

What were the results of the evaluation? After one year of CCE implementation, we found that Hindi and Math test scores of students in schools exposed to the CCE program were statistically identical to those of students who were in the comparison schools—students in the CCE schools did no better than those in the comparison schools in either subject. Hence, we can conclude that CCE program did not improve learning outcomes.

So why didn't a programme designed ostensibly to improve learning outcomes have a positive impact? There are two main reasons why programmes fail – they are either not designed to address the key problem and/or they are not implemented in the field properly. In this section, we use a combination of anecdotes and hard data to unpack what may have gone wrong.

To ensure strong implementation, the program 'suppliers' (here, the teachers) have to be well

²A pedagogy intervention which involved remedial teaching at the level of student's ability by grouping was also evaluated at the same time (TaRL). Please refer to (Banerjee 2017) which describes the findings from this and related interventions.

trained, monitored and given support when required. We found that while over 90 percent of teachers were trained on CCE, only 41 percent of teachers in primary and 21 percent of teachers in upper primary classes maintained evaluation sheets and report cards, which are critical for recording and use of evaluation information. While this is egregious, what was more concerning was that even when records were maintained, the information from evaluations was not typically used to identify low performers, to change teaching practices or to provide feedback to students. Therefore a key underlying tenet of CCE was unmet. Interestingly, while the official CCE concept note issued by the SCERT did mention identification of low performers and recommended remediation, remediation was not covered extensively during teacher training. Given the poor learning outcomes in the state, concrete remedial measures for various causes of low performance were also not recommended as the faculty at SCERT indicated that the onus of developing remedial measures rests on the teachers. The design of CCE which involved evaluating students across 20 skills and 41 sub-skills proved to be extremely onerous. The most common concern expressed by head-teachers was that CCE was extremely time consuming. While more than two-thirds of teachers surveyed indicated that they faced problems implementing CCE, less than 10 percent indicated it positively affected teaching.

Travelling in the field and speaking to teachers, we acquired a more nuanced understanding of what CCE meant to them. A significant number of teachers viewed CCE as just an increase in the number of times a student was to be evaluated, or a need to do more projects and ensuring children were encouraged in co-curricular activities. Since the 'no detention' policy was also introduced at the same time, a few teachers took it to mean 'no exams/evaluations' more broadly and questioned the need for CCE. Examining completed evaluation sheets, we found that teachers either provided comments such as 'good/fair' or did not provide any, while a few teachers indicated that they needed a lot more training on evaluating co-curricular activities and behaviour. Even teachers who had completed evaluation sheets weren't able to indicate which of the children were low performers and why, so readiness for remediation seemed a long way off.

These conversations and observations indicate that teachers hadn't internalized the philosophy of CCE which may have affected their interest and ability to implement.

Lessons learned

Though our evaluation specifically examined Haryana's CCE programme, there are more generalisable lessons that can be distilled. Haryana's CCE wasn't designed keeping in mind ground realities of student performance or teacher's motivations – it failed as it did not focus on building basic skills or entrench a mechanism of feedback.

An ideal CCE programme would not just involve going through the motions of incorporating different evaluation tools and conducting more evaluations; it would chart out a clear process by which the evaluation data can be analysed and fed-back into the teaching learning loop. It would set clear guidelines for identifying low performer students and provide insights on types of remedial measures that can be adopted. While the administration believes that teachers are best equipped to devise their own remedial programs, insights from the field indicate that teachers do not possess this skill, and therefore a state policy with such a focus is warranted. Too many parameters for evaluation lead to significant investment of time in evaluation, time which could be put to more productive use. CCE programs in other states that have similar documentation requirements would entail similar burden as Haryana's CCE program. Interestingly, a 2014 National Council for Education Research and Training (NCERT) report indicated that many states do have such requirements as part of their CCE programme.

CCE, while having a clear underlying theory, has not been found to have a significant impact on learning outcomes and is therefore unlikely to be the programme India requires. Given the situation India faces- where a significant percent of students do not possess basic learning competencies- there is a dire need for programmes that directly focus on building basic skills. Interestingly, the pioneer of CCE in India, the CBSE, seems to have recognised the pitfalls of CCE and reinstated a system of evaluation in secondary schools close to one that existed prior to the advent of CCE. Perhaps it is time for state governments to take stock and reconsider.

References:

- 1 ASER Centre. 2012. Annual Status of Education Report 2011. New Delhi: Pratham.
- 2 ———. 2017. Annual Status of Education Report (Rural) 2016. New Delhi.
- 3 Banerjee, Abhijit, Rukmini Banerji, Esther Duflo, Harini Kannan, Shobhini Mukherji, Marc Shotland, and Michael Walton. 2017. From Proof of Concept to Scalable Policies: Challenges and Solutions, with an Application. *Journal of Economic Perspectives*, Vol. 31, Number 4
- 4 Berry, James Harini Kannan, Shobhini Mukherji, and Marc Shotland. 2018 Failure of Frequent Assessment: An Evaluation of India's Continuous and Comprehensive Evaluation Program. Working Paper, J-PAL.
- 5 Black, Paul, and Dylan Wiliam. 2009. Developing the Theory of Formative Assessment. *Educational Assessment, Evaluation and Accountability (Formerly: Journal of Personnel Evaluation in Education)* 21 (1): 5.
- 6 Central Board of Secondary Education. 2009. Quarterly Bulletin of the Central Board of Secondary Education 48 (4). 2010. Continuous and Comprehensive Evaluation: Manual for Teachers, Classes VI to VIII
- 7 Government of India. 2009. The Right of Children to Free and Compulsory Education Act 2009. Gazette of India 39 (August).
- 8 Indian Express. 2017. New assessment pattern by CBSE baffles parents, schools. Retrieved on April 23rd, 2018 from <http://indianexpress.com/article/education/new-assessment-pattern-by-cbse-baffles-parents-schools-4614918/>
- 9 Sharma, Kavita. 2014. CCE Programme/Scheme of States and UTs. Department of Elementary Education, National Council of Educational Research and Training.

Harini is a Senior Research Manager and Post-Doctoral Fellow at J-PAL South Asia at IFMR. Her interest in evidence-based policy formulation influenced her decision to work with J-PAL South Asia. She is currently working in New Delhi as a Principal Investigator on a variety of evaluations in education and health. She also works with the J-PAL South Asia training team to provide customised advisory services for various partners such as Bill & Melinda Gates Foundation, USAID, and the Governments of Haryana, Tamil Nadu and Punjab. She may be contacted at harini.kannan@ifmr.ac.in