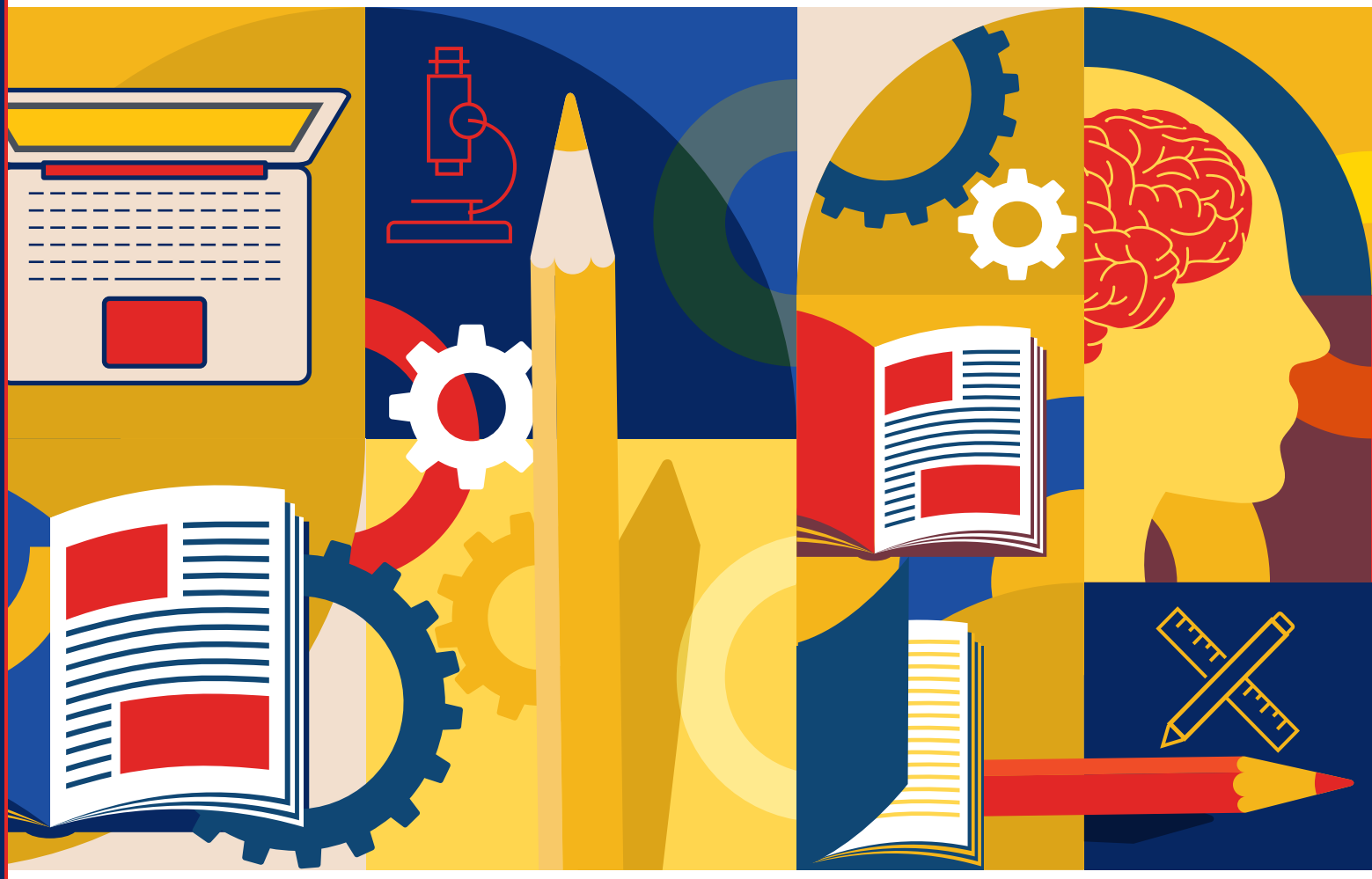
Curriculum and Learning:
Towards a Competency-based Science, Technology, Engineering, and Mathematics (STEM) Curriculum Reform in Armenia



Funded by the European Union

THE WORLD BANK
IBRD • IDA | WORLD BANK GROUP

1818 H Street NW
Washington DC 20433
Telephone: 202-473-1000
Internet: www.worldbank.org

Cover design: Alejandro Espinosa/sonideas

Contents

Graphics

Boxes

Tables

Acknowledgments

This final report has been prepared by a World Bank team led by Renata Lemos (Task Team Leader) and Maria Jose Vargas Mancera. The team includes (in alphabetical order) Melissa Adelman, Caroline Sandrine Leroy, Isil Oral Savonitto, Maria Pomes-Jimenez, Daniel Rodriguez-Segura, Anush Shahverdyan, Lucía Valdivieso, and Adriana Nikol Cordero Velazco. In earlier years, project implementation was led by Katia Marina Herrera Sosa, Hanna Katriina Alasuutari, and Anush Shahverdyan in coordination with the Center for Education Projects under the Ministry of Education, Science, Culture and Sports, the Assessment and Testing Center, the National Center for Education Development and Innovation, the National Center for Education Technology, the Republican Pedagogical-Psychological Center, and Ayb Educational Foundation.

The work was executed under Trust Fund (No TF073269) of December 28, 2018; Contract 2018/404-659; World Bank Project Number P167562 and P170001. It is part of the World Bank Group's Technical Assistance EU4Innovation STEM Pilot Activities, under the European Commission – World Bank Partnership Programme Part III for Europe and Central Asia Programmatic Single-Donor Trust Fund.

The team appreciates support and guidance from Rita Kullberg Almeida, Carolin Geginat, Narine Tadevosyan, Irina Tevosyan, and Ahmet Levent Yener at the World Bank.

Executive Summary

Around the world, basic education fails to equip many young people with the essential skills needed for employability in a 21st century job market. Almost half of students in low- and middle-income countries complete their basic education without having developed foundational literacy, numeracy, and socio-emotional skills.[1] Moreover, evolving labor markets demand 21st-century skills like collaboration and critical reasoning, which are not consistently integrated into education systems. To address these issues, it is necessary to ensure that students learn in a meaningful and consistent manner to develop more advanced skills, ultimately enabling them to become more productive citizens who can adapt to the changing nature of work.[2]

Curriculum is a potent but complex tool for improving education outcomes, with infrequent revisions of limited scope – rather than comprehensive overhauls – regularly pursued by countries. Curriculum reform is a context-specific and political process, involving many stakeholders including teachers, parents, experts, universities, and employers. It is also resource intensive, requiring technical expertise in pedagogy, individual subject matter, and multiple other areas, as well as the development and distribution of materials and trainings. Done well, curriculum reform processes usually include four key elements: *(i) a well-defined framework based on clear learning objectives; (ii) relevant, high-quality teaching and learning materials; (iii) effective and timely professional development; and (iv) informative learning assessment practices*.[3,4] Aligning these elements – such that textbook content corresponds to curricular learning standards and that those standards are assessed in exams, for example – is critical. As is engaging stakeholders at every level to ensure buy-in, monitor progress, and build capacity for the sustained implementation of reform and frequent revisions to ensure alignment with labor markets. Despite these complexities, curriculum reform is a popular policy area, with many countries engaging in minor modifications, though comprehensive overhauls as Armenia has undertaken – where content and competencies, time allocation, and pedagogy are revised across all the grades of basic education – are rare.[5]

This report documents the ambitious reimagining of Armenia's Science, Technology, Engineering, Mathematics (STEM) curriculum in basic education. It describes the main drivers of the reform's success to date, challenges faced, and rigorous evidence of positive impacts on learning from pilot implementation. With few experiences to draw from, analysis of comprehensive curriculum reforms is relatively rare, limiting the knowledge base for other countries to draw from.[6] This report aims to help fill that gap by (i) describing the reform process through the lens of the four

1 World Bank, 2022a.
2 World Bank, 2019a.
3 World Bank, 2019b.
4 Gouëdard et al., 2020.
5 For example, 23 countries in Europe reformed some part of their curriculum between 2000 and 2020 (Author's analysis of Educational Reform data from the OECD).
6 One example is Vietnam's competency-based curriculum reform in general education, launched in 2016 (Kataoka et al 2020).

main elements of effective curriculum reforms described above and (ii) evaluating the reform's implementation and impacts to date. The report focuses on activities related to STEM subjects that were funded by the *EU4Innovation Science, Technology, Engineering and Math Pilot Activities* Trust Fund. The government has carried out activities related to non-STEM subjects in parallel.

Armenia's Motivation for Reform

Armenia is striving to increase human capital for inclusive growth, innovation, and high-productivity jobs, building on significant economic growth and poverty reduction over the past two decades. While GDP growth averaged 6.2 percent per year between 2000 and 2022, Armenia still faces high rates of unemployment and inactivity, as well as a proliferation of low productivity jobs.[7] A mismatch between the formal qualifications of graduates and the skills sought by employers is an important contributor to sluggish labor market demand and overall productivity, which hamper growth and poverty reduction. Increasing human capital is therefore crucial for boosting innovation, productivity, and competitiveness, particularly as Armenia strategically competes for higher-value segments of global value chains and invests in climate change resiliency.[8,9]

Specifically, the Government of Armenia aims to equip graduates of its basic education system with high-quality, relevant skills in STEM subjects. A large-scale learning assessment, Trends in International Mathematics and Science Study (TIMSS), indicates that Armenian students' learning levels in math and science have been increasing since 2011. Armenian 4th grade students now perform on par with Europe and Central Asia (ECA) regional averages in math and slightly below them in science. However, further improvements are hampered by widely accepted challenges with the existing curriculum including fragmented, vague, and overloaded content; out-of-date learning goals compared to the pace of change in STEM disciplines and modern labor market needs; pedagogies overly reliant on lecture and limited teacher supports; and national exams misaligned with curricular goals. These are the challenges the curriculum reform tackles and aspires to turn into an opportunity for the country's further growth and development.

Armenia's STEM curriculum reform, which was initiated in 2018, aspires to transition the education system to one that is competency-based, through a focus on student-centered, inquiry-based, and outcome-oriented teaching, learning, and assessment. In particular, the objectives of the supported project are to: (i) revise and redevelop the STEM curricula for grades 1 to 12; (ii) revise and redevelop the associated pedagogical materials, including textbooks, teachers' guides, and laboratory equipment; (iii) train teachers and principals on the principles of the revised curriculum and strategies for implementation in at least one region of the country; (iv) pilot both the revised curricula and the teaching and learning materials in the same region; and (v) use conclusions drawn from the pilot program as a basis for making recommendations on the nationwide implementation of the revised curricula.

The Ministry of Education, Science, Culture and Sport (MoESCS) is leading the reform for STEM subjects in an iteratively designed and evidence-based process, to allow for learning and improvement throughout implementation. With the technical support of the World Bank and the financial support of the European Union, MoESCS utilized different types of evidence to inform the design of the reform, including (i) administrative, survey, and qualitative data to understand the Armenian education context, and (ii) expert presentations and materials to learn about international good practices in curriculum reform. Due to the COVID-19 pandemic and other unforeseen circumstances, the implementation of the curriculum began one year later than originally planned. It commenced in the 2021-2022 school year with a pilot program for grades 2, 5, 7, and 10 in schools located in the Tavush region. These four grades

7 Unemployment rates hovered around 18 percent between 2008 and 2020.
8 World Bank, 2022b.
9 World Bank, 2017b and World Bank, 2020b.

were strategically selected to provide information over the full spectrum of primary and secondary education and to allow for all subjects to be piloted. During the pilot, multiple types of quantitative and qualitative data were collected to get feedback on implementation and evaluate the impacts of the pilot. The rollout of the new curriculum in more grades and regions planned for the coming years will be based on the lessons learned from the pilot in Tavush.

Four Drivers of the Reform's Success

A well-defined framework as the basis for an effectively organized and sequenced curriculum with clear learning objectives: Committees of highly qualified national experts for each STEM subject led the design of the new curriculum utilizing a three-dimensional framework of science learning, with the support of international expert advisors. These committees used the framework from the National Academies of the USA defining Disciplinary Core Ideas, Science and Engineering Practices, and Cross-Cutting Concepts.[10] This framework provided a foundation for establishing standards, designing curricular content, and defining pedagogical approaches for each STEM subject and grade. Emphasis on progressive learning, teamwork, student-centered teaching, and real-world application helped shape the revised curriculum. The design process took approximately two years to complete, as workshops began in 2019, draft learning standards were developed and widely consulted with stakeholders throughout 2020, and final approval was given by the government for pilot implementation in 2021.

Relevant, high-quality teaching and learning materials aligned with the new curriculum: The development of new teaching and learning materials aligned with the revised curriculum followed a similar expert committee model, and these materials were provided to schools as part of the pilot. Textbook manuscripts were developed for all STEM subjects in the grades included in the Tavush pilot (grades 2,5,7, and 10), with utilization of existing textbooks and resources varying with the extent of curricular changes by subject. Other materials such as teacher guides, lab notebooks, and guidelines for project work were developed on a subject-by-subject basis. In addition, science and ICT lab infrastructure and equipment were provided to pilot schools to support their effective implementation of the new curriculum. While this aspect of the pilot faced procurement and other delays, by May 2022, nearly all schools in Tavush had been fully equipped.

Timely and effective professional development for the main implementers of curriculum – teachers, principals, and the staff who support their work: A multistakeholder team led by MoESCS designed teacher training, teacher mentoring, and administrative staff training sessions to equip all educators in the pilot region of Tavush with the knowledge and skills needed to implement the new curriculum. Prior to the first school year of implementation, teachers participated in both pedagogical and subject-specific content training in a combination of online and in-person sessions. Analysis of recorded online sessions finds that trainers were well prepared and utilized effective instructional practices, but that the online format limited participatory opportunities. Training was followed by ongoing mentoring throughout the school year to help navigate the new content, master new pedagogical approaches, and identify and address challenges as they arose. In focus group discussions at the end of the first year of implementation, teachers reported these mentoring sessions as being highly beneficial. Training sessions were also offered to school principals and administrative staff to improve their understanding of the new curriculum and strengthen their management and community engagement skills.

Informative learning assessment practices in the classroom and across the system: Efforts to align the student learning assessment system with the revised curriculum include increased use of formative assessment in the classroom and the development of a new national framework for student assessment. The pedagogical training sessions for teachers and principals aimed to equip them with various tools, techniques, and methods of formative assessment as an ongoing

10 National Academies of the USA – Next Generation Science Learning Standards https://www.nextgenscience.org/

means of monitoring student learning in the classroom. In addition, a new national student assessment framework was approved in early 2024 to guide how and when to measure what students know and can do compared to new curricular expectations, providing valuable data for teachers and principals to adjust their practices if needed and for evaluating the reform's impacts on an ongoing basis.

Impacts and Lessons from Implementation

A robust evaluation finds that the curriculum reform pilot had a positive and sizeable impact on student achievement in math and science, equivalent to an additional six months of learning over the two years of implementation. The evaluation exploits variation in the regions and grades affected by the pilot implementation following a differences-in-differences approach. The first difference is the difference in learning between consecutive pairs of grades, where one grade was not targeted by the reform and the other was. The second difference is the difference in the learning gap between pairs of grades across Tavush and comparison regions. Using the curriculum-based student assessments designed as part of the reform process, the evaluation finds that students who experienced the two-year pilot of the new curriculum made significant relative learning gains in both math and science subjects tested, equivalent to six months of additional learning. Importantly, the implementation of the curriculum reform in Tavush seems to have had a positive impact across students, teachers, and principals of different characteristics.

Analysis of pilot implementation points to three areas for continued attention and investment as the reform process continues: limiting curriculum overload, implementing the new national learning assessment framework, and continuing stakeholder engagement. Regarding curriculum design, overload issues identified in the first year of the pilot through a side-by-side analysis of the old versus new curriculum as well as teacher surveys and focus groups were promptly addressed through a careful reduction in topics and skills. Regarding learning assessment, implementing the new national student assessment framework is an important next step for achieving full implementation of the curriculum reform. The framework supports a comprehensive test development cycle applying modern psychometric methodologies considering the current international professional standard for analysis, equating, scoring, and reporting of results, which guarantee the comparability of results over time. Finally, surveys show that teachers and principals view the curriculum reform positively and feel mostly prepared for its implementation. However, many identify shortages of staff, materials, or infrastructure and excessive workloads as stumbling blocks. As the reform is rolled out in more grades and regions, continued engagement with teachers and principals, as well as parents and other stakeholders, will be crucial to maintaining support and identifying and addressing issues to enable the curriculum's effective implementation.

The successes and lessons learned from Armenia's STEM curriculum reform can help other countries set expectations for what it takes to succeed with such an ambitious policy change. First, building on local expertise and integrating stakeholder feedback from the design stage onwards take time, but these investments are required to develop enduring capacity and support for the long and complex process of curriculum reform. The working group model with national subject matter and education experts involved in the process at every stage – from design through textbook development to teacher training implementation – helped further advance already strong local capacity and ensure coherence across stages. In addition, the regular collection and use of stakeholder feedback, particularly from teachers, helped build support for the reform process, design effective capacity building, and identify and address implementation issues. **Second**, providing new teaching and learning materials requires a realistic timeline including opportunity to iterate on drafts and address inevitable delivery challenges. Some materials were too complex and had to be revised based on teacher feedback, while technically specific equipment for science labs required time-intensive procurement processes as well as unanticipated upgrades to basic enabling infrastructure, which required more time to complete. **Third**, high-quality training for teachers and administrators to implement the new curriculum is critical, but ongoing mentoring may be equally important. While there is no data yet on changes in teachers' practices, their feedback and independent assessment of recorded sessions suggests that the training for the pilot was of good quality. However, there were limited opportunities for active participation in large training sessions, and the mentoring program set up to help

and guide teachers seems to have provided the support they needed in their day-to-day work to implement the new curriculum. **Finally**, the reform process benefits from being iterative, as feedback from the pilot on all aspects from continued overload issues to the format of teacher trainings are being incorporated into future rollouts to strengthen implementation. In addition, given the systemic nature of curricular reform, needs for additional comprehensive changes in areas like learning assessment should be anticipated in planning.

1. Curriculum reform for a changing Armenia

The global context

Human capital is key to development, growth, and poverty reduction. Despite this, young adults around the world are completing basic education without the necessary skills that make them employable, productive citizens. In low- and middle-income countries, nearly half of students are completing basic education without acquiring adequate foundational literacy, numeracy, and socio-emotional skills.[11] The lack of adequate foundational and socio-emotional skills is rarely remediated once young people exit basic education and therefore often directly translates into a shortage of skills in the labor force, reducing the human capital in many countries. Moreover, twenty-first-century skills such as collaboration, critical reasoning, citizenship, and growth mindset are becoming increasingly indispensable to meet the demands of a changing labor market, but are not yet taught effectively in all education systems.[12] Changing these dynamics requires ensuring that children learn meaningfully every school year and stay in school longer to build more advanced skills, become productive citizens, and effectively adapt in an uncertain world in which the nature of jobs is changing.[13]

Curriculum can be a powerful lever for improving learning in basic education. While the causes of the global learning crisis are multifaceted, what is taught and how it is taught plays a critical role in determining the effectiveness of education systems. In the broadest terms, curriculum can be defined as an educational plan and often refers to the materials, documents, and methods used for teaching and learning. Clear, coherent, and relevant curricula are critical to effective education. They enable i) teachers to prepare and deliver high-quality instruction, ii) learning progress to be measured, and iii) interventions to improve learning—from the classroom to the policy level – to be planned, implemented, and evaluated.[14] A comprehensive examination of the Trends in International Mathematics and Science Study (TIMSS) dataset concluded that differences in aspects of the curriculum such as coherence and focus are associated to learning

11 World Bank, 2022a.
12 Partnership for 21st Century Skills.
13 World Bank, 2019a.
14 World Bank, 2019b.

gains across countries.[15] Conversely, overloaded curricula – that cover too many topics in too little time relative to the pace of student mastery – contribute to lower learning levels in basic education across developing countries.[16] A recent reform in Tanzania that focused early grade curriculum on core subjects to provide enough time for students to master foundational skills in early grades increased learning and potentially reduced later dropout.[17] Curriculum can also be a powerful lever for shaping the beliefs and attitudes of the next generation. For example, the introduction of a new political science curriculum for high school students in China impacted their political views as measured through surveys conducted when they were in college.[18]

Yet curriculum reform is an inherently context-specific and complex political process with many stakeholders. With a few exceptions, curriculum reform that successfully impacts learning is a once-in-a-generation undertaking. Because curriculum reform is neither frequent nor standardized, there is no set roadmap or blueprint for its implementation. It is one of the most sensitive and high-stakes reforms undertaken in education systems, and resistance to change is often much stronger than the desire to change. Various interest groups have a stake in curriculum change including: i) teachers and students who interact directly with the curriculum; ii) parents who have views on what their children should learn; iii) academic experts who have views on what students should learn and at what age; iv) universities who have expectations as to what students should be able to do when they graduate from school; and v) employers who have expectations as to what people should be able to do when they enter the workforce. With so many interests at play, curriculum change is often a complex political process.

Creating a well planned and executed curriculum reform is also a highly technical and resource-intensive process. Curriculum design should be informed by a sophisticated understanding of the fundamental elements of learning and teaching, and needs to address what is to be learned, how it is to be learned, and why it is to be learned. It also needs to outline how learning is evaluated and what resources are required to attain the outlined objectives. In addition, effective curriculum design involves highly technical processes such as defining new standards for: i) content and assessment, ii) teacher qualifications, iii) educational resources and learning materials, iv) management, and v) evaluation implementation. Additionally, curriculum materials such as textbooks, syllabi, or IT support must be designed and supplied to enable schools and teachers to implement the curricular reform.[19] In practice, these myriad components all need to be considered in the process, as all are consequential for the successful implementation of a curriculum.

Despite these challenges, changes in various aspects of curriculum are a popular area of policy reform in education systems around the world. In recent years, many countries have engaged in modifying aspects of their curriculum – whether it be the content taught and competencies developed, the allocation of time by topic and subject, or the pedagogical methods used to deliver the curriculum. OECD's database of education policy reforms currently shows that 119 reforms related to curriculum, qualifications, and standards took place in 23 selected education systems in Europe from 2000 to 2020. Most of these reforms – 89 in total – were carried out in the second decade, suggesting that the need for curriculum reforms is increasing across Europe (Figure 1). Curriculum changes were made across different levels and areas of education ranging from early grade reading to technical subjects in higher education. While partial curricular reforms are more common, a complete overhaul of the education curriculum like the one being undertaken in Armenia – where content and competencies, time allocation, and pedagogy are revised across all the grades of basic education – is rare, as is documented experience sharing about such comprehensive changes.[20]

15 Schmidt and Houang, 2007. Curriculum coherence is referred to the increase in depth, sophistication and complexity while bring together the fundamental, unifying ideas of the discipline across grades. TIMSS has been conducted every four years since 1995 and has been a valuable tool for monitoring international trends in mathematics and science achievement in the fourth and eighth grades. Armenia did not participate in the TIMSS 8th grade assessment in 2019 for either math or science.

16 Pritchett and Beatty, 2015.

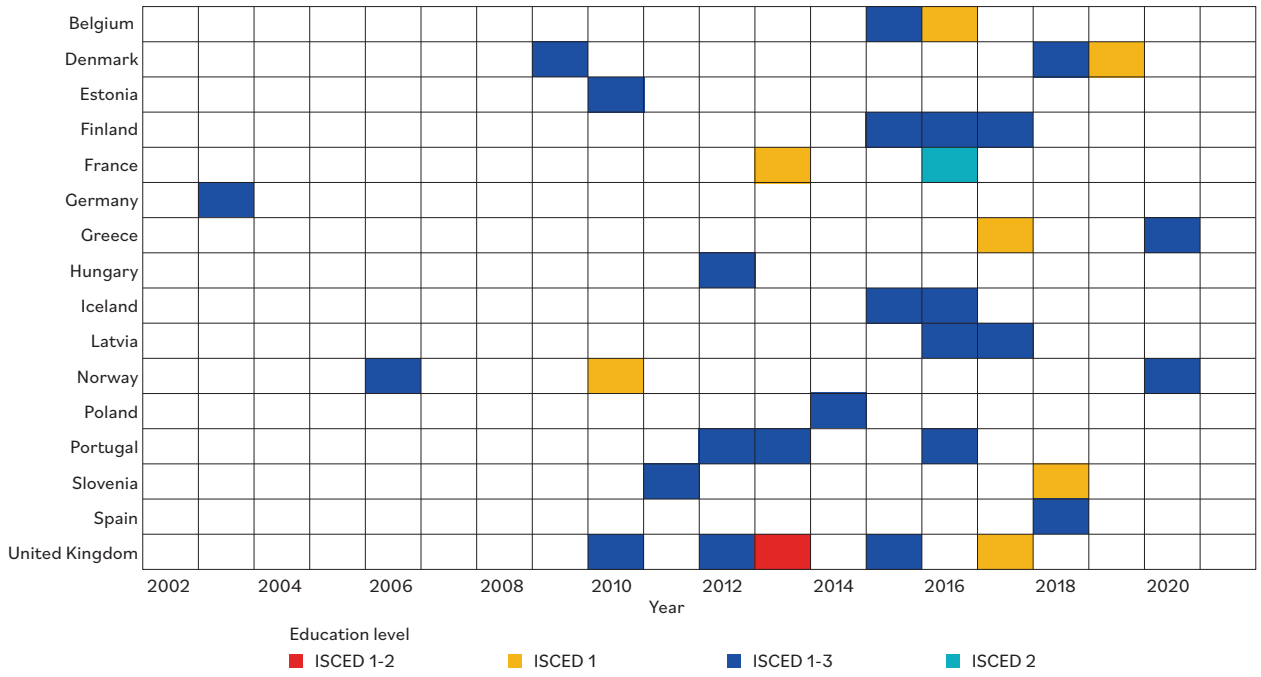
17 Rodriguez-Segura and Mbiti, 2022.

18 Cantoni et al 2017.

19 Castro Superfine, Marshall and Kelso, 2015.

20 One example is Vietnam's competency-based curriculum reform in general education, launched in 2016 (Kataoka et al 2020).

Figure 1. Many European countries have reformed parts of their basic education curriculum between 2000 and 2020

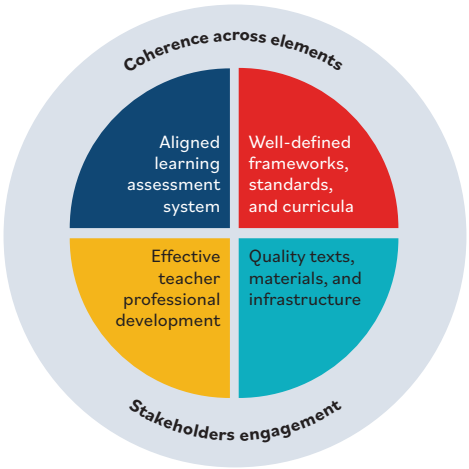


Source: Author's analysis of Educational Reform data from the OECD, data retrieved from https://www.oecd.org/education/reformsfinder.htm. ISCED 1 = Primary Education. ISCED 2 = Lower Secondary Education. ISCED 3 = Upper Secondary Education.

Global experience suggests that effective curriculum reform processes include four common elements, ensure coherence across these elements, and engage stakeholders at every level (Figure 2).[21] The first element is a well-defined framework that lays out the overall objectives for the reform with a clear vision of the knowledge, skills, and behaviors expected to be learned from the new curriculum. Beyond an overarching framework, effective reforms clearly specify new learning standards and the detailed curricula to achieve them. Standards often aspire for rigor and excellence but should also incorporate performance "floors" to ensure that all students are learning up to minimum acceptable levels. The second element is the development and provision of relevant teaching and learning materials as well as needed supporting infrastructure in the school setting such as laboratories and other learning spaces. Curricula itself and accompanying materials are also more effective when they are developed in consultation with teachers and provide clear guidance for teachers on the role of different pedagogical approaches, while keeping teachers' capacities in mind. The third element is ensuring effective and timely professional development to equip teachers with methods and strategies to teach the new materials in an optimal way to guide learning. Finally, the fourth element is to assess the learning outcomes intended by the new curriculum to help identify the strengths and weaknesses of the new system and make sure all components complement each other and lead to improved outcomes for students.[22] Without attention to coherence across all four elements, reforms are less likely to impact student learning, as curricular standards may not fully align with how teachers allot instructional time and the pedagogical approaches they use, or with the content or competencies assessed in exams.[23,24] Throughout this entire process, an important undertaking is engaging multiple stakeholders

21 World Bank, 2019b and Gouëdard et al., 2020.
22 World Bank, 2019b.
23 See Atuhurra and Kaffenberger 2022 for a quantification of incoherence in Tanzania and Uganda's education systems.
24 Rodriguez-Segura (2020) assesses the impacts of an early grade Spanish literacy curriculum reform that accompanied a first-grade social promotion policy in Costa Rica, finding that a lack of professional development and support for teachers contributed to students dropping out in subsequent grades.

Figure 2. Effective curriculum reforms align education systems and stakeholders for change



Source: Derived from World Bank, 2019b and Gouëdard et al., 2020.

from all levels of the education system – from frequent consultations with school principals and teachers to provide feedback on the development of the new curricula or materials, to consultations with students to understand how the new curricula is being taught and skills learned or whether it requires further changes – and devoting the necessary resources for institutional capacity building geared towards policy makers to optimize decision-making and support.[25]

This report describes the ambitious process undertaken to reimagine Armenia's curriculum in basic education, and documents the main successes and challenges experienced so far to contribute to the global knowledge base on curriculum reform. The report focuses on activities related to STEM subjects that were funded by the *EU4Innovation Science, Technology, Engineering and Math Pilot Activities* Trust Fund. The government has carried out activities related to non-STEM subjects in parallel. After a brief overview of the Armenian context, the report describes the reform process through the lens of the four main elements of effective curriculum reforms: i) the design of new frameworks, standards, and curricula, ii) the development and supply of new teaching and learning materials, iii) the provision of principal and teacher professional development, and iv) the alignment of the learning assessment system with the revised curriculum. The report then evaluates each of these elements using different analytical approaches and provides a rigorous estimation of the overall impact of the pilot implementation in the Tavush region on student learning. The report closes with a set of initial lessons learned from the ongoing process in Armenia that may be relevant to other countries planning for their own curriculum reforms.

Armenia's motivation for curriculum reform

Armenia is an upper middle-income country striving to enhance its human capital for sustainable and inclusive development. GDP growth averaged 6.2 percent per year between 2000 and 2022 and the poverty rate has fallen significantly from 32.1 percent in 2000 to 26.5 percent in 2021. Nonetheless, Armenia still suffers from high rates

25 Gouëdard et al., 2020.

of unemployment[26] and inactivity, as well as a proliferation of low productivity jobs. A mismatch between the formal qualifications of graduates and the skills sought by employers is an important contributor to poor labor market performance.[27] This gap leads to underutilized human capital, slower labor market demand, and sluggish overall productivity, all of which hamper growth and poverty reduction. Improving the quality and relevance of education is among the most effective ways to remedy these problems and help young Armenians become competent participants in society and contribute to their country's economy.

Firms in Armenia find it challenging to recruit and retain employees with the necessary workforce skills, creating a constraint to innovation and formal sector employment. A 2013 survey of firms shows that a third to half of Armenian firms report difficulties in recruiting professional staff, service workers, and technicians. A skills shortage was identified as the source of this obstacle particularly by firms at the technological frontier who introduced new products, invested in research and development, and upgraded their existing products during the boom years in Armenia between 1995 and 2005.[28] The same survey revealed that most employers are dissatisfied with the quality of education, especially with the provision of practical skills and up-to-date knowledge. More recently, a 2020 World Bank Enterprise Survey revealed that 11.3 percent of firms report an inadequately educated workforce to be the single most important business environment obstacle, and 27.5 percent of firms are offering formal training to their workforce.[29]

Alleviating the skill constraints of firms will be crucial in boosting innovation, productivity, and competitiveness, particularly as Armenia strategically competes for higher-value segments of global value chains based on cutting-edge technologies.[30] Since 2006, the IT and high-technology sector has become one of the fastest growing sectors in the country. According to the 2014 IT Skills Assessment in Armenia Report, the demand for IT specialists is expected to continue growing at an estimated rate of 17 percent per year, however the current number and quality of graduates is not enough to meet industry demand.[31] A skills assessment of graduates in IT and Engineering by the Enterprise Incubator Foundation revealed that 73 percent of firms find that the practical knowledge of graduates is below expectations. [32] Moreover, as technologies continue to evolve, producing basic education graduates with strong and up-to-date STEM skills is a critical upstream input to Armenia's broader strategic efforts to shift into higher-value segments of technology-driven global value chains like biotech and earthquake engineering.[33]

Boosting productivity and increasing STEM skills will be even more critical to the economic growth of Armenia in the future given its aging population and the growing global challenges of climate change. By 2040, one in four Armenians will be over 60 years old, meaning that a shrinking pool of workers will need to provide for a growing pool of elderly people. A considerable proportion of the population currently holds low-wage, low-productivity jobs in the informal sector, which in turn affects innovation, and economic growth. Without improving productivity, it will be challenging to support an increasingly elderly population, which already receives a significant share (63 percent) of social protection expenditures through pensions.[34] Moreover, Armenia is already one of the most disaster-prone countries in the Eastern Europe and Central Asia region. As weather shocks continue to become more frequent and intense, demand for STEM skills in many sectors including agricultural adaptation, natural resource management, disaster risk management, and "green economy" industries like sustainable infrastructure and low-carbon manufacturing is expected to increase.[35]

26 Unemployment rates hovered around 18 percent between 2008 and 2020.
27 World Bank, 2022b.
28 World Bank, 2017a.
29 World Bank, 2020a.
30 World Bank, 2017b and World Bank, 2020b.
31 World Bank, 2014.
32 *Ibid.*
33 World Bank, 2020b.
34 *ibid.* pp. 5-6.
35 World Bank 2019; ILO, 2019 and Kwauk, 2021

To address these challenges, the Government of Armenia aims to ensure that graduates of its basic education system have high-quality, relevant skills in STEM subjects. The government recognizes that the accelerated pace of change in science, technology, engineering, and related disciplines represents both a challenge and an opportunity for the national education system. It therefore seeks to provide students with the knowledge and skills they need to (i) understand and use science, technology, and math in their everyday lives; and (ii) pursue careers requiring additional and advanced training in STEM subjects beyond basic education.[36] The Armenia Development Strategy 2014-2025 highlighted the importance of strengthening science and technology to increase its competitive advantages and to achieve its goal of developing a knowledge-based economy. This is reflected in its strategic objectives of: (i) enhancing human capital through better access to quality services including healthcare, education, culture, and basic infrastructure and (ii) expanding employment through high-productivity and decently paid jobs. The Government Program (2021-2026) demonstrates a commitment to the development of education and science. The education activities and targets aim a creative, civilized, proactive, capable, competitive citizen, and build on the rollout of new general education standards in all classes of all schools by 2026, new textbooks and educational materials and an upgrade to school infrastructure, particularly for laboratory furniture and equipment. It highlights strengthening Armenia's competitive advantages based on science and technology as a condition for development. Finally, the Education Sector Strategy towards 2023 establishes a legal basis for the adoption of a competency-based curriculum throughout the country, followed by an action plan and costing. It highlights the adoption of a new competency-based curriculum guided by learning outcomes, with a focus on strengthening STEM subjects, green education, and foreign languages. Two of the main underlying principles behind this reform is to ensure environmental sustainability through "green" education to raise public awareness and support Armenia's transition to a green economy and to strengthen the internationalization of Armenian education.

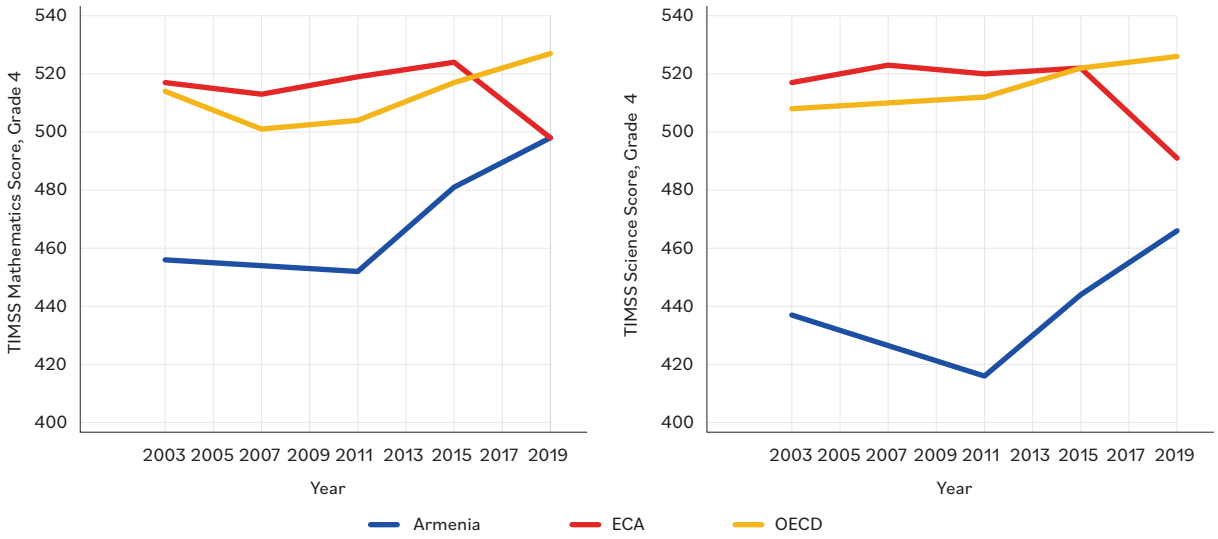
Armenian students' learning levels in math and science have been increasing since 2011, but there is still room for improvement. A large-scale learning assessment, Trends in International Mathematics and Science Study (TIMSS), indicates that Armenian 4th grade students now perform on par with ECA regional averages in math and slightly below them in science (Figure 3). In 2019, 64 percent of Armenia's students achieved the Intermediate TIMSS benchmark in 4th grade math, meaning they can apply basic math concepts in simple situations. For 4th grade science, 47 percent of Armenian students achieved this benchmark. The highest scores achieved by Armenian 4th graders in both science and math were recorded in 2019 (Figure 3).[37]

Over the last two TIMSS rounds, math scores improved particularly among low performers. Math scores for the 5 percent of students with the lowest proficiency levels increased by 72 points, from 316 in 2011 to 388 in 2019. The top performers, however, have only slightly improved. This suggests that while the "floor" has been raised, there is still much room for improvement as there are few students achieving the highest scores for math in 2019 (Figure 4). This is significant because the qualifications of students performing at a higher level is an important measure of skills relevance, future productivity, and the health and competitiveness of the labor market.[38]

Poor educational content is widely considered a major contributing factor to the performance challenges amongst students. Prior to the reform described below, the national curriculum of Armenia was fragmented, vague and overloaded, with curricular goals not closely aligned to modern labor market needs. In numerous instances, links between the subject learning objectives, requirements for learners and subject core content in all levels of general education were not clear. Curricula were overloaded with disconnected factual knowledge previously introduced and not updated or integrated with explicit learning goals. Where learning goals were included, they were articulated in vague terms and in some cases were identical for lower and upper secondary school. Pacing was also often too fast, leaving many students

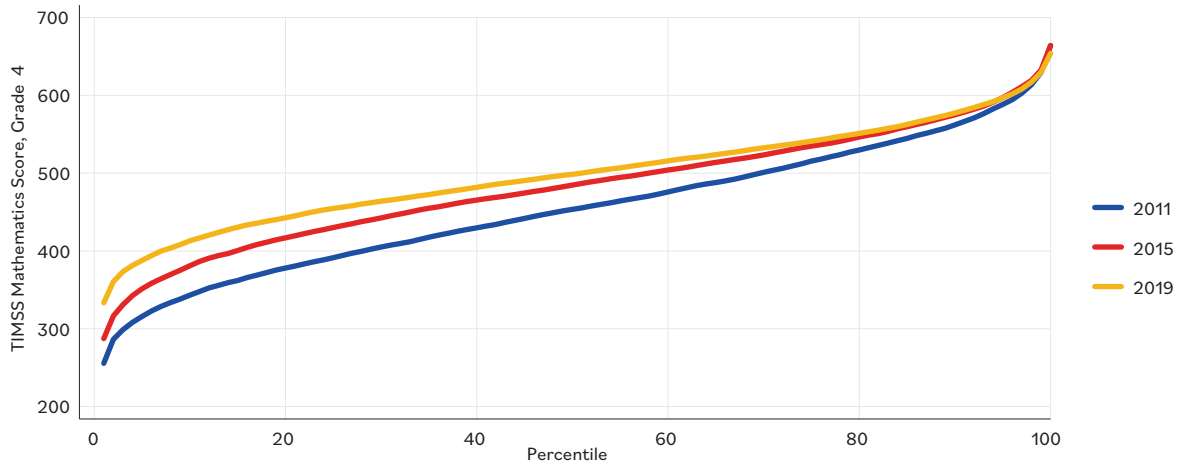
36 The Government of the Republic of Armenia, 2021.
37 Armenia did not participate in the 8th grade assessment for TIMSS 2019.
38 Science scores follow a similar pattern.

Figure 3. Armenia's TIMSS scores have improved over time but remain below OECD levels



Source: Authors' calculations using IEA data, 2003–2019.

Figure 4. Fourth grade TIMSS math score improvements have been concentrated among lower performing students, with minor changes for high performing students



Source: Authors' calculations using IEA data, 2011–2019.

without a solid foundation or substantive understanding of the topics at hand. Moreover, beyond the curriculum itself, the teaching practices, instructional materials, and textbooks in Armenia were outdated and at times ineffective.[39]

The STEM subjects' curriculum suffered from several important deficiencies regarding relevance, pedagogy, and assessment of learning outcomes. With respect to relevance, the STEM curriculum was out-of-date in some important

39 World Bank, 2016.

ways. First, it did not fully reflect the range and breadth of advances in science and math at the frontiers of knowledge that have repercussions for what students should be taught as they gain foundational skills. Second, the local and global challenges to which scientific and technological knowledge can be applied continue to change and evolve. With respect to pedagogy, the curriculum relied too heavily on lecture-based instruction and lacked student-centered and active-learning techniques. Both the amount and quality of guidance available to teachers to improve their pedagogy required updating and revision. Finally, with respect to assessment of learning outcomes, the curriculum lacked alignment with national exams. More specifically, the system was reliant on end-of-cycle high-stakes national assessments that emphasized memorization or specific content proficiency at the expense of depth of understanding, ability to apply knowledge to real-world situations, and general- and domain-specific critical thinking abilities.

Education has long been a policy priority in Armenia, with important improvements over the past two decades leading up to the current reform. In 2005, Armenia became a signatory to the Bologna Declaration, an agreement signed by 45 countries to make their academic systems more compatible with European standards. As a result, Armenia switched to a 12-year system of education beginning with the 2006-2007 academic year. The move replaced the traditional 10-year system of the Soviet era with a new three-tier structure comprising four years of elementary school, five years of middle school, and three years of high school. This was followed by efforts to improve certain aspects of the education system into the following decade, including extending preschool coverage and promoting greater links between higher education institutions and the labor market. These efforts culminated in the need to reform the basic education curriculum.

Overview of the Armenian reform

Armenia is currently implementing an ambitious reform of the basic education curriculum to help students reach their full potential and drive the social, economic, and political development of the country. The decision to reform the education curriculum was taken by the new government following the Velvet Revolution in 2018. One of the goals of this reform is to transition Armenia's education system into one that is competency-based by focusing on inquiry-based, student-centered, and outcome-oriented teaching, learning, and assessment. This is being accomplished by overhauling the entire curriculum at all levels of basic education. In 2023-2024, the curriculum reform has entered the third year of the pilot stage and results will inform the eventual national rollout.

The MoESCS is leading the reform process for STEM subjects with the technical support of the World Bank and the financial support of the European Union. The reform is an opportunity for the government to incorporate new priorities throughout all basic education subjects. Although the STEM subjects are the focus of the Bank support, all school subjects including humanities and social sciences are undergoing revision and redevelopment. The shared principles of the reform that are being incorporated across subjects are: (i) inclusive education, that is, ensuring a diversity of viewpoints including those of minority groups and traditionally excluded groups; (ii) gender equality and the promotion of attitudes and behaviors that reflect equality and respect; and (iii) sustainable development.

The reform focuses on fully modernizing the curricula and training teachers on up-to-date, student-centered pedagogies. In particular, the objectives of the reform are to: (i) revise and redevelop the STEM curricula for grades 1 to 12; (ii) revise and redevelop the associated pedagogical materials, including textbooks, teachers' guides, selected exams or examination blueprints, and laboratory equipment; (iii) train teachers on the principles of the revised curriculum and strategies for implementation in at least one region of the country; (iv) pilot both the revised curricula and the teaching and learning materials in the same region; and (v) use conclusions drawn from the pilot program as a basis for making recommendations on the nationwide implementation of the revised curricula. The specific steps taken in the reform design process are outlined in Box 1. These steps were created by local and international experts in a collaborative effort.

Box 1. Armenia's reform was purposefully designed to be an iterative process[40]

The Armenian Curriculum Reform	Step 1: Detailing the new curriculum and new pedagogical approaches as well as learning outcomes, resulting in documents on subject curriculum standards.
	Step 2: Establishing detailed blueprints for each subject and each grade as the basis for constructing tests for a summative assessment.
	Step 3: Developing new teaching and learning materials for the subjects and grades to be piloted in the Tavush region.
	Step 4: Providing STEM and ICT laboratory equipment and furniture to schools in the Tavush region
	Step 5: Providing training to the teachers involved in the Tavush pilot.
	Step 6: Providing regular monitoring and mentoring for the teachers in the Tavush pilot.
	Step 7: Strengthening the scale-up capacity of the National Center for Education Development and Innovation (NCEDI)
	Step 8: Developing new teaching and learning materials for other grades, beginning with pilot grades for the 2022/2023 academic year in Tavush.
	Step 9: Getting feedback from the teachers during and after the pilot in Tavush.
	Step 10: Using feedback from the first pilot year to revise the curriculum and teaching and learning materials.
	Step 11: Using the first two years of the reform to evaluate its effectiveness

Source: Derived from the EU4Innovation STEM Pilot Activities.

The reform is being implemented gradually to learn, iterate, and strengthen the new curriculum. At every juncture of design and implementation, a comprehensive array of qualitative and quantitative evidence has been systematically gathered and analyzed. The specific evidence gathered is detailed in Box 2. This evidence has not only guided adjustments to the processes underpinning curriculum revision but has also helped the MoESCS in making well-informed decisions to strengthen its pilot phase and ensure a strong nationwide implementation.

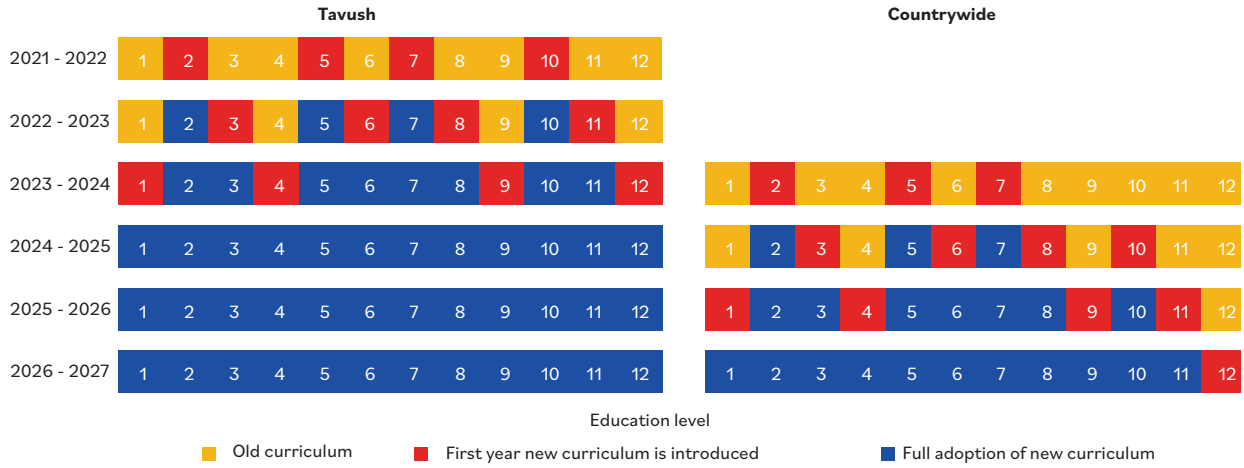
MoESCS chose to start the curriculum update in the 2021-2022 school year, piloting the curriculum for grades 2, 5, 7 and 10 in schools in the Tavush region (Figure 5, Figure 6). These four grades were strategically selected to provide information over the full spectrum of primary and secondary education and to allow for all subjects to be piloted. The choice to include grade 2 rather than

Figure 5. The new curriculum started its piloting in the Tavush region

Source: Derived from the EU4Innovation STEM Pilot Activities.

40 See Annex 1 for a detailed list.

Figure 6. The new curriculum will gradually roll out to more grades and regions every year



Source: Derived from the EU4Innovation STEM Pilot Activities and National Center for Education Development and Innovation (NCEDI).

grade 1 was taken because children in grade 2 are already adapted to the school setting. A valuable lesson that can be learned by conducting the pilot over a three-year span is to identify whether there are 'connection' problems, that is, to see if children who followed the old curriculum face difficulties when introduced to the new curriculum and to see if these difficulties persist as they progress through the education system.

It is important to acknowledge that there were initial setbacks that caused delays to the curriculum revisions and piloting. In early 2020, project activities were adjusted to comply with COVID-19 global safety precautions, which forced workshops, teacher assessments, and other activities to move online or be delayed. In mid-to-late 2020, Armenia experienced unrest due to conflict in Nagorno-Karabakh, causing many activities to temporarily halt. Additionally, the public review process to advise on the new reform was started in mid-2020 and feedback was received over the course of several months. Given the importance of reflecting on and incorporating this feedback, the initial timeline for the reform was adjusted. These multiple challenges contributed to pushing out the beginning of the pilot by one school year.

Box 2. An evidence-driven reform: How different types of evidence informed the design and implementation of the curriculum reform in Armenia

The process of curriculum reform in Armenia is particularly notable for its prioritization of evidence-based decision making, utilizing data of many different types – existing and new, qualitative and quantitative, local and global – to inform the design, implementation, and revision of the pilot reform. Local officials and experts in government and multiple organizations, international advisors, and the World Bank team worked together to compile national and international evidence to support choices made to reform the curricula. This evidence can be summarized in four categories, based on its specific purpose as shown below:[41]

Data was collected to understand the Armenian context:

Administrative data from the MoESCS and teachers survey data from TIMSS to understand teacher profile and policies. Administrative data from the National Center for Education Technology (NaCET) Statistical Database and teacher surveys collected during Armenia's participation in TIMSS and TALIS were used to perform a descriptive analysis of teacher profiles in Armenia, including demographics, educational background and tenure, professional development and training, salary structure and teaching expectations, challenges and practices, collaborations and relationships, job satisfaction and motivation.

Classroom observations to understand teacher practices in the classroom. The World Bank's freely available classroom observation tool *Teach* was implemented to understand how instruction is organized in Armenian classrooms and to inform the design of teacher training. *Teach* measures (i) the time teachers spend on learning and the extent to which students are on task, and (ii) the quality of teaching practices that help develop students' socio-emotional and cognitive skills. The *Teach* modules were implemented in a selection of 20 primary schools in two regions of Armenia: Tavush and Yerevan.[42] Classroom observations captured the practices of 80 teachers in grades 2, 4, 5, 7 and 8 across several subjects, including geography, math, and natural science.

Qualitative inputs from education experts to support collaborative revisions and develop context-specific toolkits and training. Information was collected through numerous different activities with officials and experts in the Armenian education system to support the collaborative revision of project materials and curricula structure, workshops with subject specific content experts, and the creation of toolkits based on teacher needs.[43]

Survey data on teacher instructional skills and professional development needs. The National Center for Education Technology (NaCET) and the World Bank Armenia education team designed and implemented a teacher survey in 2019 that was utilized to inform the design of the new curriculum and the modules of the teacher training. The survey was conducted in all schools across Armenia by NaCET in December 2019 and received over 35,000 responses. Teacher and school characteristics were collected, and teachers were asked about their instructional skills and experience teaching, their professional development needs, salary structure, and use of instructional materials and classroom practices.

Information was put together to learn about international good practices and what works in curriculum reform:

International expert presentations and best practice materials such as blueprints and high-quality textbooks to inform revisions of the curriculum. During several workshops with local experts held between 2019 and 2022, international experts

41 These categories are not mutually exclusive in that some data fed into several categories such as the teacher survey, focus group findings and the results of implementing the *Teach* tool in Armenia.

42 It is important to note that this study was implemented during the COVID-19 pandemic when Armenian schools were operating at limited capacity, with classes reduced to half their normal size.

43 This information was collected by Ayb Educational Foundation, which was competitively commissioned to undertake the teacher training for the pilot.

compiled and provided evidence-based materials that described international good practices to support revisions to the content, new pedagogical approaches, and learning outcomes for all the subjects in the new curricula; to discuss the best ways to approach assessments in the new curricula; and to develop new teaching and learning materials.

Data was collected to get feedback on the pilot implementation in Tavush:

Focus group interviews with teachers to gather feedback on the piloting of the revised curriculum. At the end of the 2021-2022 school year, 19 focus group interviews were conducted with grades 2, 5, 7, and 10 teachers who piloted the new curriculum to obtain information on successes and challenges they experienced during the pilot. The questions discussed included teachers' views on the new curriculum, new methods of teaching, how student learning was impacted by the new curriculum, how new learning assessment structures worked, what experiences and challenges they had with the new teaching and learning materials, how mentoring affected the pilot, and how collaboration culture played into the delivery of the new curriculum.

Local mentor observation journals to gather information on perceptions about training implementation. During and after the teacher training, local mentors were asked to reflect on how teachers were implementing the new curriculum. The mentors were trained on the feedback process: various journals were required to be filled in for each type of interaction with the teachers to track their progress and activities. The journals covered fields such as success or achievement of teachers, issues and challenges voiced by the teachers, solutions discussed, and specific steps taken. Local mentors reported pedagogical, technical, and logistical challenges in their observation journals.

Data was collected to evaluate the pilot implementation in Tavush:

*Teacher survey to track information on time spent in the classroom and classroom activity alignment with the revised curriculum to assess whether the curriculum was overloaded***.** The survey covered 260 mathematics and 162 science teachers in pilot schools and was conducted in May 2022. It asked about the time used for each subject to deliver instruction on the new curriculum as well as what the teachers thought about the time allocated to different topics in the curriculum.

Curriculum-based student learning assessments developed and carried out to evaluate the effect of the reform on students' understanding and proficiency in selected subjects and grades for the 2021-2022 and 2022-2023 school years. Assessment design and development included the analysis of the content and structure of the old and new curriculum standards, item development and item mapping from international assessments released items. 7,095 students in 2022 and 9,981 students in 2023 were assessed in math and one additional STEM subject area according to their grade. Students were assessed in pilot grades and subsequent non-pilot grades in Tavush and non-pilot Shirak and Lori regions and Yerevan for comparison.

Student, teachers, and school principal surveys to better understand the context under which the curriculum reform was taking place. Student surveys gathered information on students' demographics, household assets, educational background, attitudes towards learning, perceptions of teacher practices and socioemotional skills. Similarly, teacher surveys gathered information on teachers' demographics, teaching methods, perspectives on curriculum reform, perceptions of school management practices, and monitoring and evaluation techniques. Finally, school principal surveys gathered information on their demographics, management strategies, and their perceptions of curriculum reform.

2. Designing the new curriculum

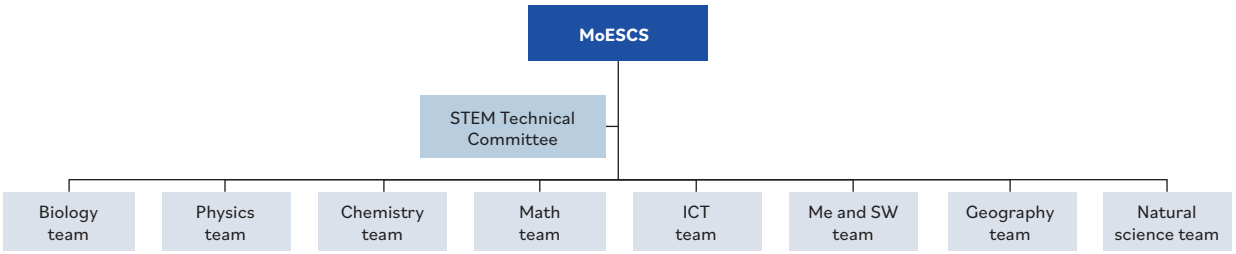
The human capital behind the design of the new curriculum

A working group approach was used in the curriculum reform process to (i) leverage national and international expertise and (ii) ensure coherence across all elements of the reform. MoESCS was responsible for overall coordination and implementation, while the World Bank team provided technical support in setting up the management structure and identifying the products to be delivered. Two tiers of management and administration were set up, each with explicit roles and responsibilities: (i) the STEM Technical Committee (TC), and (ii) the subject specific sub-committees (Figure 7). The STEM TC was comprised of (i) MoESCS representatives who provided overall guidance and direction, (ii) local STEM experts (teachers and scholars) who provided content related guidance, and (iii) international experts with experience in curriculum reform, mathematics and science, textbooks, teaching and learning materials, and assessment development, who provided global good practice guidelines and case studies to be followed. The committee created the pedagogical principles and outlined the content that needed to be taught in math and science education. It also oversaw the production of syllabi, lesson plans, and teacher guides for math and science subjects for each grade (grades 1-12) that were created by the sub-committees. Keeping the same working groups involved across each component ensured consistency and reduced "learning leakage" in that experience and knowledge from previous steps were not lost with those experts being replaced with new ones who were not familiar with the preceding activities of the reform effort.

Highly qualified national experts were selected to participate in subject-specific committees to lead the new curriculum's design for each STEM subject. The selection process included an open call for applicants in each of the subject areas of specialization.[44] There were a total of 260 candidates, of whom 42 were ultimately selected, with between 13 to 25 percent of candidates being selected to participate in each subject-specific committee (Figure 8). On average, selected candidates were better qualified than the general pool of applicants. All applicants for the subject expert positions had a university degree, mostly in their teaching subject. The selected candidates also had significantly more post-graduate studies: 76 percent held master's degrees and 55 percent held PhDs in comparison to a respective 51 percent and 35 percent, amongst the entire pool of applicants. The selected candidates brought important experience into the design process. According to their CVs, 98 percent of the selected candidates had experience in subject matters, 95 percent showed professional achievement credentials, 86 percent had experience as a teacher, 48 percent as a University Instructor, 52 percent had published a scientific article, and 36 percent had direct experience with curriculum design.

44 Me and the Surrounding World, Geography, Biology, Science, DLCS, Mathematics, Assessment, Chemistry and Physics.

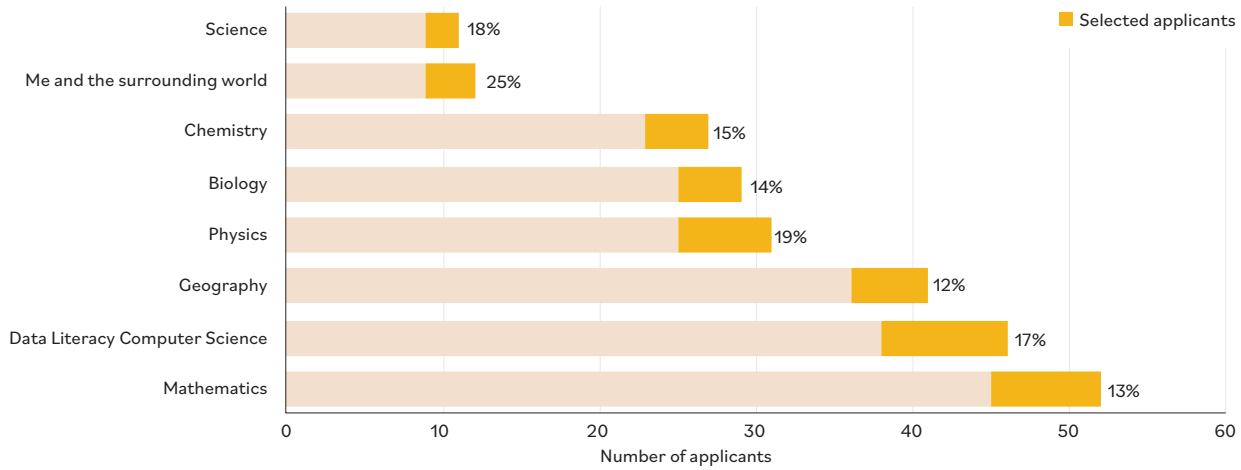
Figure 7. The working group structure of the reform utilized local and global expertise



Source: Derived from the EU4Innovation STEM Pilot Activities.

Note: The STEM Technical Committee is comprised of representatives of the MoESCS, local education experts and international education experts. The specific subject teams were led by local subject specialists, who received regular support from international subject experts based on their needs. The subject teams created the new subject standards and new assessment blueprints for the revised curriculum documents.

Figure 8. Local expertise was well-utilized as membership in sub-committees was competitive for all subjects



Source: Authors' calculations using coded data from documents submitted by applicants.

The subject specific sub-committees consisted of Armenian experts for each of the STEM subjects: physics, chemistry, biology, mathematics, computer science and digital literacy, geography, 'me and the surrounding world' (grades 1-4), nature (grades 5, 6), natural science (grades 10, 11) and geography.

This expert-led model attempted to balance the need for cutting-edge knowledge with building national capacity for sustainability and institutionalization of reforms. International experts introduced European science education principles and frameworks at the global frontier of STEM education and national experts helped ensure the integration and contextualization of these ideas into Armenia's new basic curricular principles, all under the leadership and coordination of MoESCS. At the same time, national and private organizations played important roles in implementing and monitoring the reform, including the National Center for Education Development and Innovation (NCEDI), the Republican Pedagogical-Psychological Center, and the Ayb Educational Foundation.

The defining framework for curricular changes

The reform was guided by a three-dimensional framework of science learning from the National Academies of the USA: Disciplinary Core Ideas, Science and Engineering Practices, and Cross Cutting Concepts.[45] Disciplinary Core Ideas (DCIs) are fundamental ideas necessary for understanding a given subject discipline. For example, in the physical sciences in grade 2, a DCI is that "heating or cooling a substance may cause changes that can be observed. Sometimes these changes are reversible, and sometimes they are not."[46] Science and Engineering Practices describe behaviors that make up scientific inquiry and the work of engineers in design and building, for example in grade 2 the practice of "engaging in argument from evidence". Cross Cutting Concepts enable the identification of connections across the domains of science such as patterns, cause and effect, structure and function, and the influence of engineering, technology, and science on society and the natural world. Together, these three elements of the science learning framework define standards, or performance expectations, for each subject and grade. For example, the above elements come together to define a grade 2 physical science standard that students should be able to "construct an argument with evidence that some changes caused by heating or cooling can be reversed and some cannot".

The national subject-specific sub-committees utilized this framework to establish new standards and curricular content for each STEM subject and grade. The DCIs that underpin each subject were established by the working groups and then used to guide the definition of the content, pedagogical approaches, and expected learning outcomes for all the subjects in the new curriculum. New pedagogical approaches emphasizing teamwork, student-centered teaching, project-based learning, learning from real life situations, and communication were introduced throughout the curriculum. The resulting revised curriculum is intended to be internally coherent, building progressively from grade 1 to grade 12.

Consequently, the structure and presentation of the curriculum were significantly changed, most evidently by maintaining at all levels down to the topic taught in each grade the reference to its corresponding learning outcome and DCI. For example, the grade 2 mathematics for the old curriculum included an overall teaching and learning objective for the whole grade, that was then articulated in learning objectives for each of the four topics taught, each accompanied by a section specifying the corresponding result (Table 1). In the pilot version of the new curriculum, internal coherence is emphasized by identifying and evidencing the connection for each topic and grade to learning outcomes at the core standard level as emerging from the DCIs (Table 2). For example, in the old curriculum, mathematics learners in grade 2 would be taught numbers below 100, with the stated topic objective of "knowing the names and writing forms of integers between 20 and 100", in line with the overall teaching and learning objective for the grade of "study the two-digit numbers". Under the new curriculum, these learners would be looking at numbers within 100 with the purpose of developing the skills to read, write and compare the two-digit numbers, and building the skills to perform operations with them. The learning outcome for the topic is that the student can recognize, read and write the two-digit numbers and its connection to the learning outcomes emerging from the DCIs for grades 1 through 4 is made specific. Namely, these are numbers and number systems (DCI Level 1), number sets (DCI Level 2) and "read and write multi-digit numbers, know the order and class composition of numbers" (Learning outcome S1) (Table 3). This detailed structure and internal coherence is maintained for all grades and topics.

45 National Academies of the USA – Next Generation Science Learning Standards https://www.nextgenscience.org/
46 Examples taken from DCI Arrangements of the Next Generation Science Standards

Table 1. An example from the old curriculum: Grade 2 Mathematics

Overall teaching and learning objective for Grade 2 Mathematics
Study the two-digit and three-digit numbers above 20, their comparison, operations of addition and subtraction within 1000; give idea about the multiplication and division operations; teach the multiplication table for one-digit numbers; continue introducing the geometric images, shapes, values, their measurement units; develop the ability to formulate a question.
Main teaching and learning objective for Topic 1 Two-digit numbers
• Teach the names and writing forms of numbers below 100, their comparison, addition and subtraction operations; • Build and develop the skill to compare the numbers; • Use the numbers to count the quantity of objects in a group and to enumerate those; • Use the numbers and values to solve elementary applied problems; • Teach how to use simple tables.
Study of this topic will give the learners opportunity to:
• Know the names and writing forms of integers between 20 and 100, the order composition of two-digit numbers; • Be able to count forward within 100 and backward in 1s, 2s, 5s, 10; • Collect data (also non-numerical) through inquiries and note them down; • Listen to, read and comprehend a problem composed of two-three simple sentences; • Participate in discussions, use others' data, participate in team work.

Table 2. An example from the new curriculum: Grade 2 Mathematics

Overall teaching and learning objective for Grade 2 Mathematics – Two Digit numbers
Study the two-digit and three-digit numbers above 20, their comparison, operations of addition and subtraction within 1000; give idea about the multiplication and division operations; teach the multiplication table for one-digit numbers; continue introducing the geometric images, shapes, values, their measurement units; develop the ability to formulate a question.
Purpose
• Build and develop the skills to read, write and compare the two-digit numbers • Build and develop skills to perform operations with the two-digit numbers, to add and subtract in column • Introduce and apply the concept of approximation
Learning Outcomes
• Recognize, read and write the two-digit and three-digit numbers and present them in the form of the sum of ordinal additives. • Count forward within 100 and backward in 1s, 2s, (4, 6 ,8...), in 5s (15, 20, 25), and in 10s (80, 70, 60,...). • Add and subtract within 100, also orally. • Apply the commutative of addition and associativity of addition. • Approximate the two-digit numbers to the nearest tens. • Compare the numbers within 100. • Organize the numbers in ascending and descending order. • Guess and continue the simple patterns. • Find the unknown components of the arithmetic operations. • Understand, reproduce the assumption and claim of the problem. • Solve one or two step problems, applying various arithmetic operations (less by ..../more by ...., total, and so forth).

Content
• Numbers within 100, round tens • Writing form and composition of two-¬digit numbers (point, decimal point) • Read and write the two-digit numbers in numerals and letters • Present the two-digit numbers in form of sum of ordinal additives • Compare the two-digit numbers • Approximate the two-digit numbers • Add and subtract one-digit and two-digit numbers • Add and subtract the round ten to a digital number • Add and subtract two-digit numbers, without changing the order of numbers • Add and subtract two-digit number with change of the order of numbers • Add and subtract two-digit numbers in column • Solve problems
Link with the learning outcomes of the Core Standard
S1, S5, S6, S7, S8, S9, S12, S16, S23, S25, S29, S35

Table 3. An example for the connecting between the new curriculum and DCI: Grade 2 mathematics

Disciplinary core ideas		
Level 1	**Level 2**	**Learning Outcomes**
Numbers, number systems	Numbers, sets	1. Read and write multi-digit numbers, know the order and class composition of numbers
		2. Recognize the proper fraction.
	Arithmetic, algebraic expressions and operations	1. Read a numerical expression with 2-3 operations; write the sentence with 2-3 arithmetic operations in form of a numerical expression.
		2. Perform arithmetic operations with multi-digit numbers.
		3. Calculate the value of a numerical expression, using the laws and performance order of the arithmetic operations (also in brackets).
		4. Know the components of arithmetic operations. Determine the unknown component in the arithmetic operations.
		5. Find the given part of a number or the number, based on the given part.
		6. Find the unknown value in the time-speed-distance relation.
	Comparing the numbers	1. Compare and arrange the numbers by ascending or descending order.
		2. Compare the fractions with the same denominator, and the same numerator.

The development of state and subject standards

Between 2019 and 2020, workshops and feedback sessions were held by the STEM TC to provide subject-specific sub-committees with an overview of educational reform based on international good practices to support the development of standards and curricula – to work on the content, the new pedagogical approaches, and the learning outcomes for all the subjects in the new curricula, to discuss the best ways to approach assessments in the new curricula, and to develop new teaching and learning materials, and to engage stakeholders. For example, an initial 4-day workshop was in the first semester of 2019, where education experts, teachers and leaders, university professors, and the broader education community in Armenia were invited to discuss education curriculum and its ways forward in Armenia. The workshop held interactive sessions on standards, student learning outcomes, subject areas, distribution of hours and assessment for the new STEM curriculum. A second workshop focused on the development of subject standards was held in the second semester of 2019. The objective was to support the national working groups in developing the science and mathematics subject standards. The workshop also engaged with local providers of education services to leverage national expertise as much as possible.[47] A third workshop focused on curriculum revisions, covering item blueprints and student assessments was held in the first semester of 2020. The aim of the workshop was to introduce to the expert-teachers the different methods and techniques on student assessments targeting the summative and formative types, with a final goal to develop assessment blueprints with proper conventional numbering per each STEM curriculum.

The State Standards of General Education, encompassing all the suggested revisions to the learning standards and the curriculum, were approved by the Government of Armenia in February 2021 and new subject learning standards were adopted for pilot implementation in the Tavush region in April 2021. National subject-specific sub-committees finalized the draft learning standards for all subjects by May 2020. The STEM TC provided feedback on the draft learning standards developed. Several rounds of consultations took place during the second semester of 2020: (i) in July 2020, consultations with international experts were held to discuss progress and feedback on learning standards, (ii) in August and September 2020, public consultations were held in Armenia to receive feedback on the draft learning standards, (iii) a follow-up consultation was help in November 2020 with international experts once feedback from public discussions were incorporated into the draft learning standards to discuss changes and steps forwards on finalizing the learning standards and learning materials.

47 For example, TUMO Center for Creative Technologies, a well-known private provider of free-of-charge extracurricular education programs for adolescents, participated by facilitating a presentation on ICT in education.

3. Developing teaching and learning materials

Teaching and learning materials are essential tools. They make lessons more interesting, practical, and appealing, enabling teachers to effectively convey concepts and improve students' performance. These materials can also foster active participation, skill acquisition, and self-confidence among students. By utilizing instructional materials, teachers can create a stimulating learning environment that supports students' achievement and overall educational experience. Research has shown that teachers with access to off-the-shelf high-quality instruction materials increased math achievement, and gains are higher for students of low performing teachers.[48] The revision and update of teaching and learning materials played an important role in Armenia's curriculum reform.

Between 2020 and 2021, additional workshops and feedback sessions were held by the STEM TC to provide subject-specific sub-committees with international best practices on the development of teaching and learning materials and their coherence with other documents guiding the curriculum reform. A fourth workshop on Learning Materials was held virtually in the first semester of 2020. The workshop introduced the local expert working groups to the development of learning materials and to the links between standards, syllabi, and textbooks. A fifth workshop on the development of teaching and learning materials catered to a general audience of learning professionals in the STEM and Humanities sectors was held in the first semester of 2021. A sixth workshop covered these same topics but was specifically tailored to Math and other STEM professionals. These two workshops covered topics on design, layout, publishing, and intellectual property.

The creation and revision of teaching and learning materials was a collective effort. MoESCS led the process and joined monthly meetings to ensure the teaching and learning materials were in line with the country's education policy. Local expert groups were assembled for each subject, including the experts who served on sub-committees. World Bank experts were also involved and supported the local experts with advice and guidance upon request. Three main types of teaching and learning materials were revised and provided for the pilot in Tavush schools. These include: (i) printed student learning materials such as draft textbooks and workbooks, (ii) printed teacher guides and aides, and (iii) updated science and ICT laboratory equipment and furniture.

48 Jackson and Makarin, 2018.

The development of teacher guides and student textbooks and workbooks

A group of local experts, including many of the subject-specific sub-committee members, were recruited to develop a set of grade-specific textbook manuscripts for all STEM subjects in the curriculum pilot. Each subject was assigned a chief editor who oversaw the development of the manuscripts for all grades included in the pilot (grades 2, 5, 7, and 10).[49] This ensured that the textbook content was aligned with the revised learning goals and that there was appropriate progression between grades. The chief editor led a team of experts and worked in close coordination with other chief editors to ensure consistency across subject matter. The chief editor determined the number of writers required for his or her respective subject, based on the level of complexity of the content and anticipated length of the manuscript.

The approach to textbook manuscript development varied by subject and grade, based on existing materials and the extent of curriculum changes.[50] Many subject manuscripts – including those for physics, biology, geography, chemistry, and mathematics – contained topics carried over from the previous edition of the textbook because the material fit the new curriculum's needs and helped achieve learning outcomes. However, these topics were updated with new instructional methods. In digital and computer literacy, for example, the main approach was to develop most of the materials from scratch since the existing textbooks were outdated. In science, the materials for grades 10 and 11 were also designed from scratch. In nature, the approach was different. The previous textbooks were based on several disciplines (physics, geography, biology, and chemistry) and each section covered that specific content, while the newly developed textbooks integrate all subjects. For example, the topic of "water" is introduced both from the point of view of chemistry and geography. Nonetheless, every textbook manuscript sought to answer the following key questions: (i) what are the students expected to do once they have completed this manuscript?, (ii) what questions will students be expected to answer?, (iii) what tasks will students be expected to perform?, (iv) what techniques and skills will students be expected to have acquired?, (v) what changes in behavior and attitude will be expected of the students?

For all subjects, teacher training modules, teacher guides, and other materials were also developed. These include blueprints, class plans, resource maps, and guidelines for project work. In addition, depending on the specific needs of each subject, different types of supplementary teaching and learning materials were developed from scratch. For example, a new lab notebook was developed for chemistry, and for biology, presentation files were developed to help teachers introduce new topics to students. During the first year of the pilot, the production and distribution process of materials was under the responsibility of the Republican Pedagogical-Psychological Center while the second year of the pilot was organized and implemented by NCEDI. All the printed materials were transferred to primary school No. 1 in Ijevan (Tavush region) and were subsequently distributed to all other schools based on acceptance-handover acts.

The provision of science and ICT equipment and furniture

School furniture, lab furniture, and lab equipment were also supplied to all pilot schools in Tavush. In 2020-2021, the MoESCS together with national experts developed lists and technical specifications of Science and ICT laboratory equipment and furniture to implement the revised curriculum, conducted a market study to determine the estimated cost of the provision of laboratory equipment for natural science subjects (chemistry, biology, physics, and geography) and ICT, and launched tenders and awarded contracts for laboratory equipment and furniture. Initial procurement included (i) laboratory equipment and supplies for physics, chemistry-biology, geography, and ICT, and (ii) computer equipment for STEM classrooms. One important step of the process of setting up science labs in schools was a needs assessment done through site visits to all pilot schools in Tavush, which revealed the need for additional investments.

49 Subsequent years focused on the development of manuscripts for all remaining grades.
50 Once the curriculum was officially adopted, the actual textbook development process took approximately one year, from curriculum evaluation to printing.

First, the available equipment in Tavush schools was outdated and incompatible with other new equipment to be procured, meaning additional equipment had to be procured for the project to be able to establish laboratories for the pilot in Tavush. Second, the assessment also revealed that 18 schools had no running water supply and required additional investments. In addition, minor civil works for the provision of adequate supply of water and drainage systems in chemistry and biology laboratories were carried out in 76 schools. Delays in the procurement process and these additional investments led to a delay in the provision of science and ICT equipment and furniture. However, by May 2022, ninety-six percent of the 80 Tavush middle and high schools (77 schools) had been fully equipped with laboratory furniture and infrastructure.

4. Preparing principals and teachers to adopt the new curriculum

The human capital behind the implementation of the new curriculum

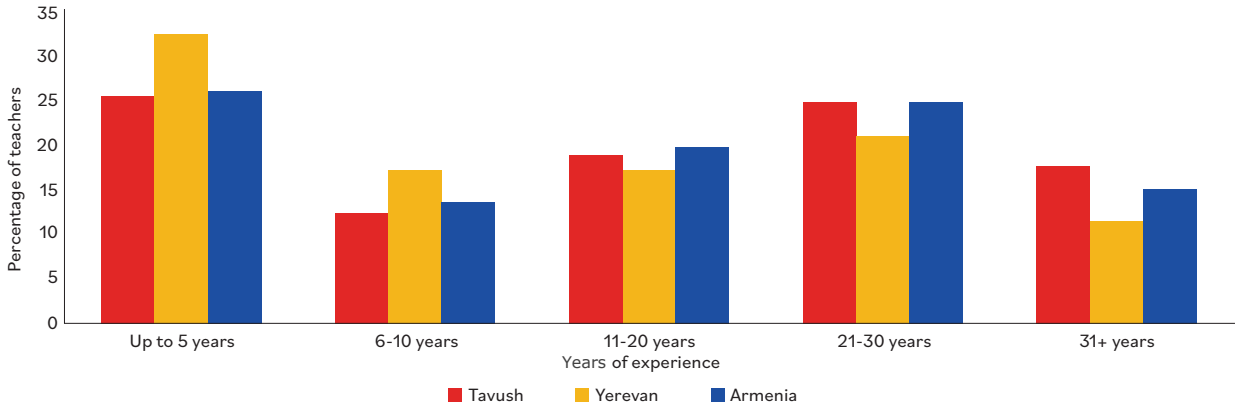
Teachers are the main implementers of the new curriculum, so equipping and supporting them effectively is critical to the success of the reform. A growing body of research indicates that teachers are the most important school-based determinant of student learning.[51] Professional development for teachers and the administrative staff who enable their work was therefore a critical part of Armenia's reform process, with training and mentoring programs developed for STEM teachers and implemented in the pilot region of Tavush. The development of the programs was informed by data on the characteristics of teachers, self-reports of their own professional practices, and independent observation data.

The Armenian teaching workforce in basic education is relatively older and experienced. Armenia is among the countries in ECA with the highest share of older teachers. Teachers are on average 47 years old, and about half are above the age of fifty.[52] In the pilot region of Tavush, almost half of the teacher population (46 percent) are older than fifty, while only 13 percent are under thirty. While only approximately 17 percent of Armenian teachers studied at the post-secondary level to become a teacher, about 75 percent have taught for over 5 years.[53] The range of teacher experience in Tavush is similar to national averages, while Yerevan has relatively younger and newer teachers (Figure 9).

Teacher self-reports and independent observations identify a range of strengths and weaknesses that are relevant to the effective implementation of the new STEM curriculum.[54] Strengths include being highly collaborative with peers

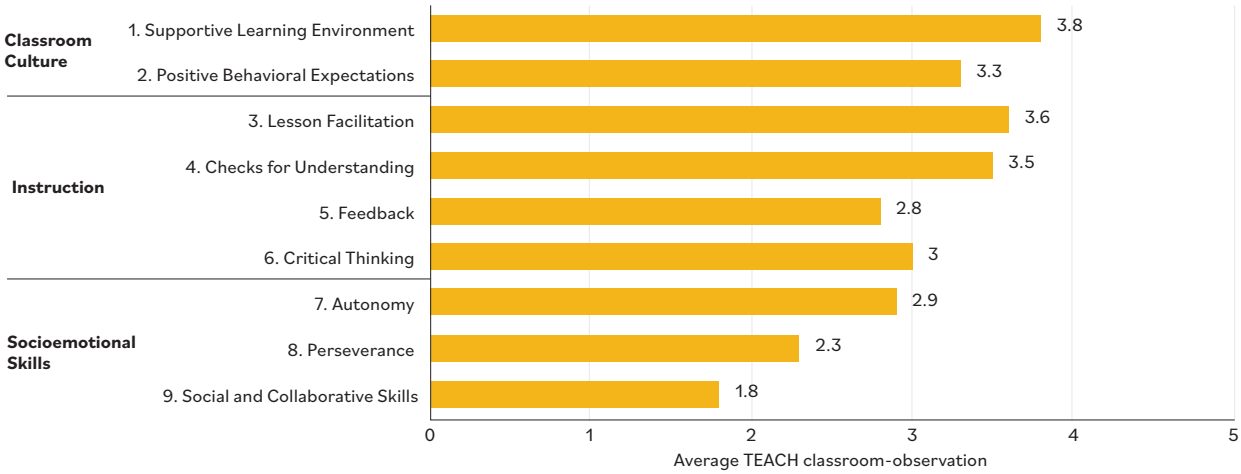
51 World Bank, 2018. For example, students with effective teachers advance 1.5 grade levels compared students with ineffective teachers who advance only 0.5 grade levels during a school year in the US (Hanushek, 1992; Rockoff, 2004).
52 World Bank, 2021.
53 World Bank, 2021.
54 World Bank, 2021.

Figure 9. About 75 percent of teachers in Armenia have been teaching for 6+ years



Source: World Bank, 2021.

Figure 10. Armenian teachers utilize many effective practices, but can improve on providing feedback, encouraging critical thinking, and promoting socioemotional skills

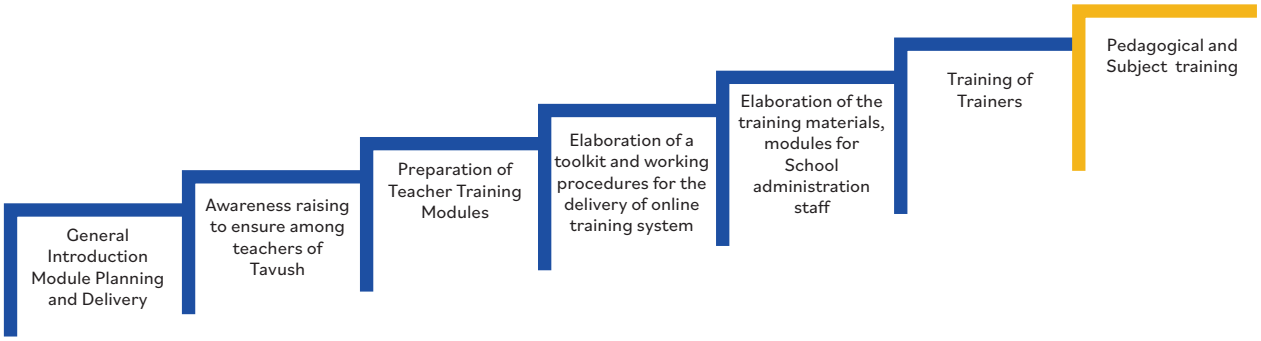


Source: World Bank, 2022c. *Note:* The TEACH tool provides a description for the different behavior levels to determine if the teacher practices could be categorized as low (L), medium (M), or high (H). After assigning quality ranges to the behaviors, the element score should be decided upon according to the overall quality of each element. Scores range from 1 to 5, with 1 being the lowest score and 5 the highest. For example, if the behavior quality ranges for autonomy are L, H and M, the overall score for that element would most likely be around a 3.

as well as school management and having a high degree of self-confidence in their ability to implement the curriculum.[55] In addition, classroom observations show that Armenian teachers have strong ability in classroom culture and instruction, with strong skills in areas such as creating a supportive learning environment, setting positive behavioral expectations, facilitating the lesson, and checking for understanding. Areas for improvement were found in providing feedback, encouraging students to think critically, promoting student autonomy, and fostering perseverance and social and collaborative skills (Figure 10).

55 World Bank, 2021.

Figure 11. The process of planning and implementing teacher training prioritized stakeholder awareness raising and feedback (May-August, 2021)



Source: Derived from the EU4Innovation STEM Pilot Activities.

Teachers lack professional development opportunities critical for continuous improvement and for effectively implementing changes such as the new curriculum. Ongoing professional development opportunities are an important mechanism for strengthening teachers' capacities, however more than half of teachers reported not having participated in any professional development in the last five years (between 2016 and 2020). Around 5 percent of teachers in Armenia report not having received any professional development since 2010.[56]

A multistakeholder team was engaged in designing the professional development activities – teacher training, teacher mentoring, and administrative staff training – for the pilot rollout of the curriculum in Tavush. The team comprised MoESCS as the overall leader of the activities, the subject experts who were engaged in the new curricula development process, NCEDI, and the Ayb Educational Foundation.[57] The design of the training materials considered data on the Tavush teaching workforce as mentioned above, incorporated stakeholder feedback (Figure 11), and included both in-person and online approaches.

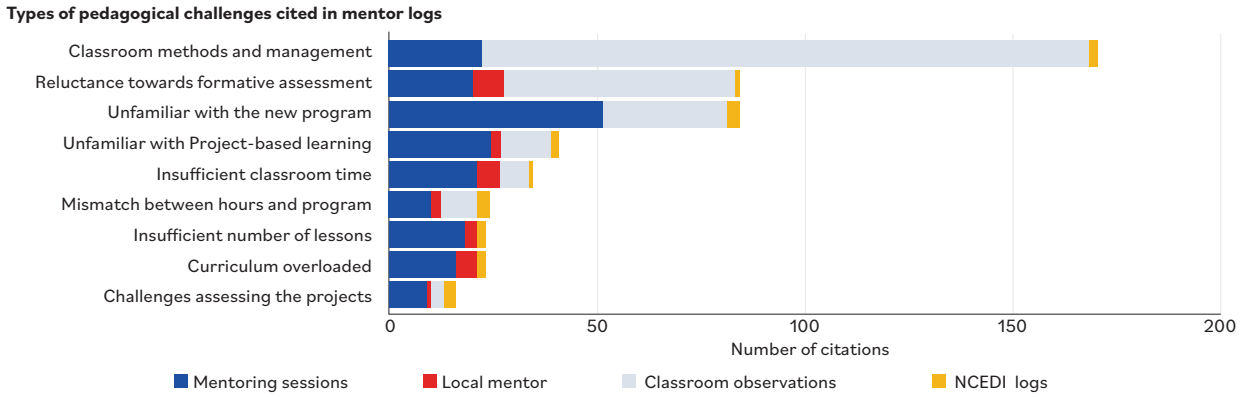
Teacher training and mentoring

The structure of the professional development program reflects several features that are considered essential to effective in-service training for teachers.[58] This included the clear and subject-specific focus as well as face-to-face sessions followed by continued mentoring and the provision of additional useful materials such as lesson plans.[59]

Pedagogical and subject-specific training sessions for teachers in the pilot grades were held before the start of the 2021-22 school year in a combination of online and in-person sessions. Teachers received pedagogical training in a combination of three online sessions, each lasting two hours, and five in-person sessions, each lasting two hours, for a total of 16 hours of pedagogical training. The three online sessions discussed professionalism, simultaneous planning of learning, and autonomous learning. The five in-person sessions covered the following topics: (i) feedback (ii) project-based learning (iii) assessment for learning (iv) task design, and (v) questioning. In terms of subject specific training, all teachers attended two hours of online training, followed by in-person training, which varied in duration

56 World Bank, 2021.
57 Ayb Educational Foundation was competitively selected as the responsible firm for implementing the training and mentoring activities.
58 Popova et al, (2022).
59 These sessions were, however, not linked to incentives such as promotions and salary implications.

Figure 12. Classroom management was the pedagogical implementation challenge most cited in mentor logs during pilot implementation of new curriculum



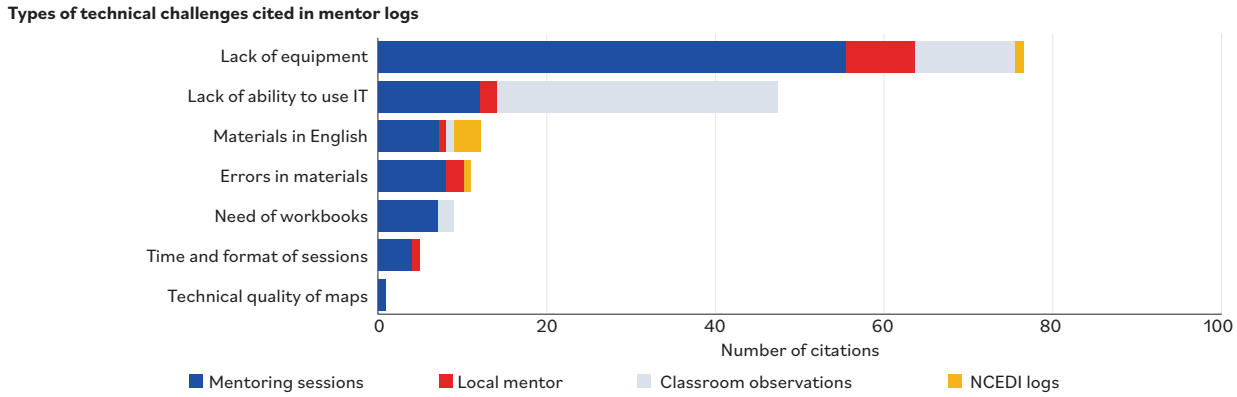
Source: AYB monitoring logs for mentoring sessions

by subject and grade ranging from 4.5 to 16 hours. Subject-specific training sessions were led by the subject experts who were part of the curriculum design working groups, and they presented the changes, underlying logic, objectives, and contents of the new curriculum.

Teacher training continued throughout the year in the form of mentoring, with subject leads becoming lead mentors and additional mentors recruited locally to act as facilitators. There was a competitive call for volunteer local mentors with specific responsibilities including: (i) explaining new topics to teachers; (ii) working with teachers on the pedagogical methodology of given topics; (iii) gathering feedback on strengths and weaknesses of training techniques; (iv) observing teachers in the classroom and providing feedback; and (v) helping teachers with student project assignments and implementation. Between November and December 2022, 27 selected local mentors participated in a series of two-hour sessions covering the fundamentals of mentorship, lesson observation, feedback provision, teaching competencies, and project-based learning. Upon finishing their training, these local mentors were paired with subject mentors and tasked with supporting them with: (i) lesson observations, (ii) small group discussions amongst teachers, and (iii) weekly subject meetings to keep the process alive and interactive. Local mentors documented their activities in logs that captured successes and challenges, analyzed below.

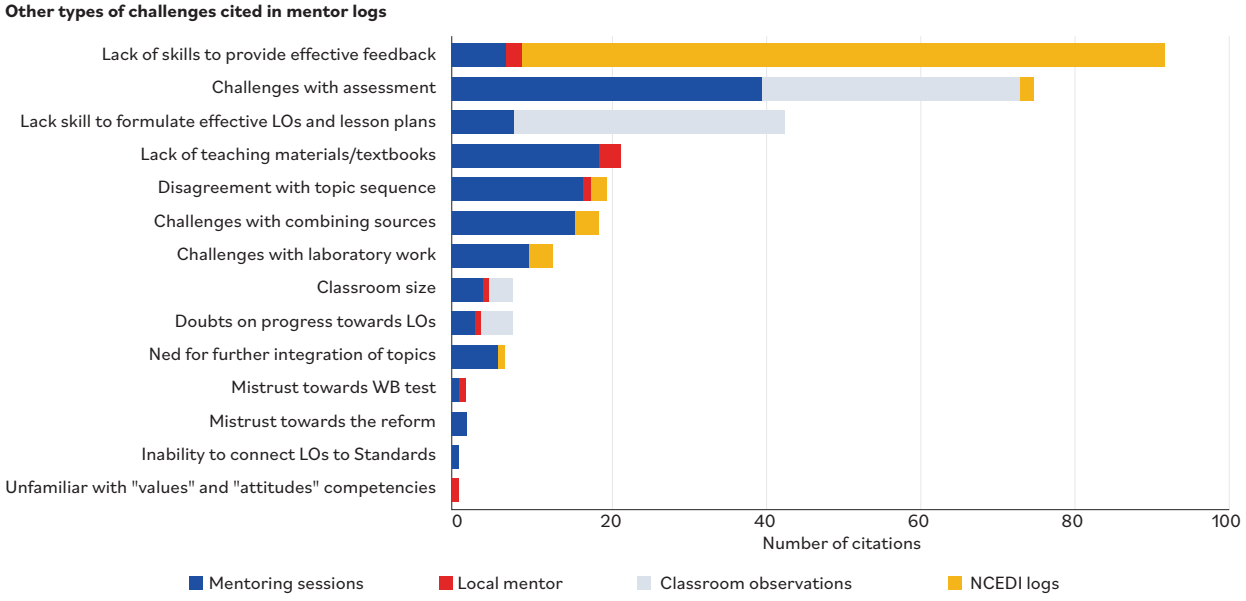
Local mentors reported pedagogical and technical and logistical challenges in their classroom observation logs. Detailed logs were kept throughout the mentoring sessions, including those collected by local mentors, NCEDI and during classroom observations. For all of these, classroom management proved to be the most frequent challenge logged and observed (Figure 12). Conversely, the unfamiliarity with the new program, a topic often discussed as part of the regular mentoring sessions was then less often perceived to be an issue once classrooms were being observed. Technical and logistical challenges raised during the mentoring sessions highlighted the absence of the necessary equipment to teach (Figure 13). Lastly, the lack of skills to provide feedback and assessment were often also registered both in the mentoring sessions and classroom observation logs (Figure 14).

Figure 13. Lack of equipment and lack of familiarity with IT were the most cited technical challenges during pilot implementation of new curriculum



Source: AYB monitoring logs of mentoring sessions.

Figure 14. Other types of frequently cited challenges during the pilot implementation of the new curriculum include lack of skills to provide feedback and to do assessment



Sources: AYB monitoring logs of mentoring sessions.

School management and administrative staff training

School principals and administrative staff were also trained for the new curriculum's implementation before the pilot started. The training aimed to effectively guide principals and deputies in the reform process and ensure a quality education for students in their schools. Assignments and in-school projects were given to participants to reinforce what they had learned and to help them develop practical methods for improving their leadership skills.

Training was offered to the school principals and administrative staff of all 82 schools in Tavush and was planned in three phases of training between August 2021 and May 2022. The first phase took place in August 2021. Five two-hour sessions were given to 145 staff members on various topics, including: (i) strategy and vision development for improving school performance; (ii) teaching and learning guidance; (iii) leading system change; (iv) development of cooperation with parents and community; and (v) team leadership, effective communication, and self-development. This was followed by five online discussions, each lasting one-and-a-half hours, reviewing previously shared material of the relevant topics from the previous phase. The second phase of this training was held in February 2022 and included ten hours of training for 103 professionals. In this phase, school principals and administrative staff presented school development projects, which were followed by panel discussions including "How to Make Your School Standout" and "Education Management Priorities in the Context of General Education Reform." Assessment and leadership were also covered in these sessions.

5. Assessing student skills under the new curriculum

Student learning assessments are key to understand the skills children are acquiring at school and to help teachers identify how they can support student learning needs. These can be formative or summative, depending on their primary purpose. The goal of formative assessments is to monitor student learning to provide ongoing feedback that can improve teaching and learning. Formative assessments are low stakes, meaning that they do not entail consequences for students, teachers, or schools. Examples of formative assessments include quizzes, asking students to draw a concept map in class to represent their understanding of a topic, or to submit one or two sentences identifying the main point of a lecture or turn in a research proposal for early feedback. All this allows teachers to gather information on student's understanding of the content and use this to adjust future lessons. On the other hand, the goal of summative assessments is to evaluate student learning at the end of an instructional unit by comparing it against some achievement standard or benchmark. Summative assessments are often high stakes, meaning that they have consequences for students, schools, or teachers. Examples of summative assessments include examinations, a final exam for approving a subject or national assessments that link results with financing. Both types of assessments provide relevant information on the teaching and learning processes but serve different purposes.[60]

Learning assessments should be closely aligned with curriculum,[61] both in terms of how their content aligns with curricular standards, as well as how assessment methods align with the pedagogical approaches prioritized in the curriculum. The aims of learning assessments are: (i) to evaluate what students learn and can do, (ii) verify students' skills and competencies, (iii) evaluate teachers' capacity to teach the curriculum, and (iv) monitor and report to education stakeholders to improve accountability. It is important to ensure that the assessment methods and tools are: (i) valid in that they measure what they propose to measure, (ii) reliable in that they measure students' skills and competencies consistently, (iii) fair in that they provide equal access to learning and assessment opportunities by different groups, and (iv) useful in that they inform students, teachers, parents and schools about students' strengths and weaknesses.

60 OECD, 2013.
61 World Bank, 2018.

Strengthening classroom formative assessments

Before the curriculum revision process started in Armenia, assessment in Armenian schools concentrated on summative as opposed to formative assessment. Armenia's education evaluation system included national and international large-scale student assessments. At the national level, state graduation examinations at the end of grades 4, 9 and 12 certify the completion of the different education levels. A Unified Examination for school leavers determines admission to higher education. Since 2003, Armenia has also participated in the international assessment TIMSS. The transition from a fully summative means of assessment to the addition of a formative assessment system was a new process in Armenia, and a process that required a fundamental change inside the classrooms.

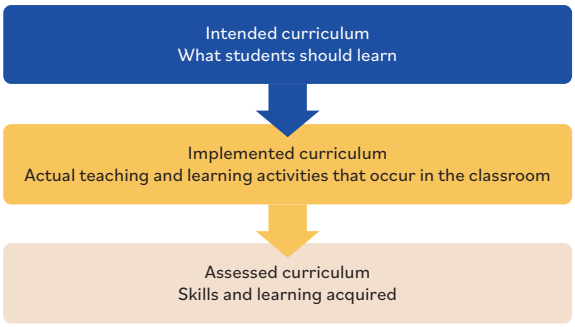
Increased use of formative assessments and a structured approach to yearly summative assessments at the school level are both part of the Armenian curriculum reform. Formative assessment was included in the pedagogical training sessions for teachers and principals in the summer of 2021, during which various tools, techniques, and methods of formative assessment were presented to the participants. The discussions during the training revealed the need for school leaders to have a full understanding of formative assessment, oversee its practice in their schools, and communicate its use properly to parents and other stakeholders, and for teachers to feel comfortable and confident in their ability to use formative assessments in the classroom. A workshop for teachers and school leaders that focused on assessment methods was organized to encourage the use of formative assessment techniques in classrooms. Several methods of formative assessment were discussed to convey the ways in which a teacher can provide continuous feedback to individual students in the classroom. In addition to this, assessment blueprints were developed by international experts during the design of the curriculum reform and the development of the teaching and learning materials. These blueprints aimed to help teachers to design and apply classroom assessments based on the new curriculum learning standards. A structured approach to summative assessment was also discussed. This approach builds from the new learning outcomes and covers how teachers think questions should be asked in formal tests or examinations, and the various levels of depth these questions should cover.

Measuring new skills in a standardized way

Aiming to assess whether the intended skills were being acquired by students in the classrooms, a tailored set of student assessments was developed. Assessing students' skills acquisition is central to monitor and evaluate the impact of curriculum reforms. This involves an examination of three interconnected components: the intended curriculum, the implemented curriculum, and the assessed curriculum (Figure 15). The intended curriculum, serving as the theoretical framework, delineates the overarching learning objectives for students across various grades and subjects. In contrast, the implemented curriculum delves into the practical application of these objectives within classrooms. The final component, the assessed curriculum, provides a tangible measure of students' skills and understanding through diverse assessment methods. To measure what students were effectively learning inside the classrooms during the piloting of the new curriculum, summative curriculum-based assessments for selected subjects for pilot grades were developed: (i) mathematics for grades 2 and 7, "Me and the Surrounding World" for grade 2, and geography for grade 7 for the 2021-2022 school year, (ii) mathematics for grades 3 and 8 and, "Me and the Surrounding World" for grade 3 and physics for grade 8 for the 2022-2023 school year. These assessments aimed to understand what test takers knew and could do based on the curricular expectations.

The design and development of the assessments was carried out by a team of international and local experts. An international team of competitively selected psychometrists and item developers was tasked with the design and development of the assessments. The experts worked in parallel and served as peer reviewers of each other's analysis given their subject specific expertise. Assessments were then revised by local experts, who provided guidance and feedback. Finally, the items were adapted and translated to Armenian.

Figure 15. The intended, implemented and assessed curriculum represent different levels of curriculum design and implementation



Source: Adapted from OECD (2013).

The process for developing the assessments included (i) a curriculum analysis to identify the overlap between the old and new curriculums, (ii) developing content specifications, which identify the objectives and skills which are to be measured by the assessment; and (iii) drafting and adapting assessment items aligned to the content specifications. The development of the assessments used evidence-centered design,[62] an approach used to ensure that the assessment aligns with its intended purpose and provides evidence of what students know and can do. This approach allows the systematically unpacking of curriculum performance expectations into multiple components. The experts identified the knowledge, skills, and abilities (KSAs) that the assessments aimed to measure by defining the content domain and learning objectives taking the overlap of old and new curriculums as the basis of the analysis. The item design process involved: (i) identifying target performance expectations, the comprehensive statements of the learning outcomes; (ii) unpacking performance dimensions: gathering information about how knowledge and skills are acquired and used in the domains, including developing a set of assessment boundaries for content knowledge, detailing the range of content and Depth of Knowledge (DOK) levels and unpacking the subject practices by defining the core aspects of the subject practices, and specifying the evidence of KSA associated with the practices; (iii) developing an integrated dimension map (content specification) to describe the essential relationships and link them to aspects of the subject practices.; and (iv) specifying design patterns to develop items that are aligned with the target learning performances. The design patterns articulate item features that are needed to elicit evidence of student proficiency.

Complementary surveys were also carried out with students, teachers, and principals to better understand the context and aid in the interpretation of the results observed in the learning assessments. Student surveys were developed and conducted to gather information on students' demographics, household assets, educational background, attitudes towards learning, perceptions of teacher practices and socioemotional skills. Similarly, teacher surveys were conducted across pilot grades capturing data on teachers' demographics, teaching methods, perspectives on curriculum reform, perceptions of school management practices, and monitoring and evaluation techniques. These surveys were applied to both STEM and non-STEM teachers. Finally, a school principal survey was also applied to gather information on their demographics, management strategies, and their perceptions of curriculum reform.

The development of the curriculum-based assessments and standardized surveys using validated questions served as a first step to consolidate an item bank to continue measuring learning and understanding student, teacher, and principal characteristics and engagement with new curriculum. Following best practices in learning assessments, these included items with different levels of complexity and competencies (i.e. knowledge, comprehension, application, analysis). Over 174 items were developed for use in the 2021-2022 school year, out of which over 50 high quality items were used as anchor items for the assessment used in the 2022-2023 school year to ensure comparability. For assessments in the 2022-2023 school year, anchor items were complemented with released items of TIMSS 2007 and TIMSS 2011 mapped to the learning objectives of the curriculum. Student, teacher, and principal questionnaires also made use of validated items and constructs found in surveys in a range of international assessments such as PISA, TIMSS and PIRLS. These were then translated to Armenian. These learning assessments and surveys were carried out in pilot grades and subsequent non-pilot grades in Tavush and non-pilot Shirak and Lori regions and Yerevan for comparison when evaluating the impact of the curriculum reform.

62 Mislevy & Haertel, 2006

Aligning the student assessment framework with the new curriculum

The education sector in Armenia has heavily relied on infrequent international assessments to measure national education outcomes. It is crucial to measure learning outcomes and use learning data to inform the implementation of reforms, especially in the context of curriculum reform. Simply adopting a new competency-based curriculum that emphasizes active learning and creative thinking will not bring about significant change on its own. The effects of curriculum reform can be weakened if the examination system remains unreformed and creates misaligned incentives for students.[63] Currently, the student exit examinations are not comparable from year to year, making it difficult to track the assessment process. The development of these examinations relies heavily on experts' opinions without a standardized approach and evidence from data analysis, threatening the reliability and validity of the tests. TIMSS has been used to provide data on 4th grade learning outcomes and benchmark to other countries, but it is conducted every four years only.

Various studies suggest that national and subnational learning assessments provide insights at the system level that cannot be obtained through classroom assessments by teachers. Policy makers need to understand whether students are mastering the national curriculum, identify areas where students are stronger or weaker, determine if certain population groups are lagging and by how much, and identify factors associated with better student achievement. It is not possible to aggregate the results of classroom-level formative assessments by teachers to obtain reliable system-level information. National assessments across the country can play a crucial role in tracking systemwide progress and can help identify cases where trends or levels of student achievement differ between subnational assessments, thus ensuring the quality of assessments at different levels.[64] Once classroom and national assessments are established, participating in regional or global assessments that enable performance benchmarking with other countries can be highly beneficial. The goal is to develop assessment systems that are aligned internally but serve different needs.

Workshops were organized to discuss how to enhance the MoESCS's capacity to design and implement a robust evaluation system. The workshops brought together assessment specialists and psychometricians from the Assessment and Testing Center (ATC) and the NCEDI. Three main topics were covered: (i) Item Response Theory (IRT): This statistical model is used internationally for creating valid and reliable large-scale assessments. The current approach used by the ATC relies on Classical Test Theory (CTT). The workshop highlighted the advantages of using IRT for test development, data analysis, and scoring of national test data. It emphasized the importance of pre-testing items to ensure content and psychometric properties, as well as the need for a different test architecture to allow comparability over time. The use of IRT provides precise information for every item and student, facilitating valid and reliable tests and informing teaching methodologies and educational policy decisions. (ii) Comparison of Statistical Models: The workshop compared the results obtained from IRT analysis and CTT analysis of previous National University Entrance tests in Mathematics and Physics. The comparison showcased the differences between the two models and revealed item quality issues that were not detected using CTT. IRT analysis also provided insights into the reliability of the test and the areas where decisions on test scores are made, which were not possible using CTT. (iii) The test development cycle: The workshop presented the typical test development cycle used by international assessment bodies. It outlined the necessary timeframes, staff requirements, and training needed for effective implementation using IRT methodology. The involvement of various teacher-groups in the test development process was emphasized, along with the advantages of a test development cycle that maintains test quality and comparability of results over time. Overall, the workshops aimed to improve the assessment system in Armenia – by introducing IRT methodology, addressing item quality issues, and implementing a comprehensive test development cycle.

63 World Bank, 2018.
64 World Bank, 2018.

MoESCS has recently taken a significant step by approving a new National Student Assessment Framework that will serve as a comprehensive guide for the development and implementation of assessments aimed at evaluating the education system. It specifies the types of assessments, their intended objectives, the process of their development and construction, the procedures for their administration, the methodology for analyzing the results, as well as the utilization and dissemination of the assessment outcomes. The framework includes three types of assessments: internal assessments, international assessments that are already in use, and external standardized assessments that need further development. This comprehensive approach ensures that all aspects of the education system are evaluated effectively. One of the most significant aspects of this new framework is the commitment from the MoESCS to introduce modern psychometric methodologies, such as IRT, for the development of large-scale tests for external standardized assessments. This demonstrates the Ministry's dedication to utilizing advanced techniques to ensure the accuracy and reliability of the assessments. If implemented well, the external standardized assessments described in the National Student Assessment Framework will have a substantial positive impact at all levels, including students, schools, and regional and national stakeholders. The detailed and precise information provided by these assessments will be essential for tracking the results of curriculum implementation over time and across different populations. This will enable the MoESCS to make evidence-based decisions related to the curriculum in the future, ensuring continuous improvement and progress in the education system.

6. Evaluating the implementation of the new curriculum

Assessing curricular changes

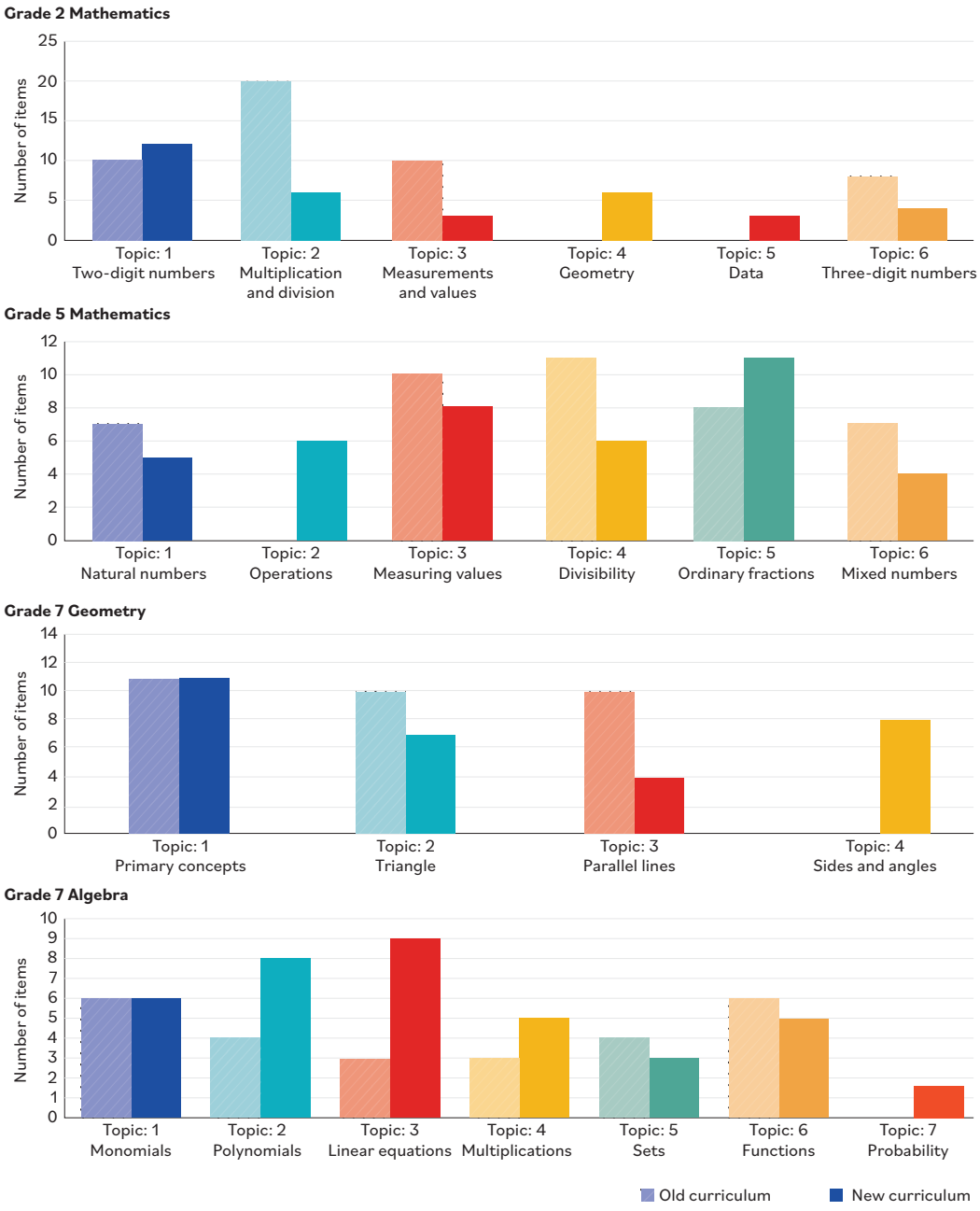
An analysis of curricular changes was undertaken to assess whether the curriculum reform unintentionally led to curriculum expansion, content overload, or perceived overload. Curriculum expansion refers to the inclusion of new topics without appropriately considering old topics to be removed. Content overload refers to a disproportionate amount of content to be taught compared to time available for teaching. Perceived overload refers to how teachers experience the delivery of the curriculum and its related materials in the classroom.[65] To understand whether this was the case in the Armenian curriculum reform, three exercises were done. First, a side-by-side analysis of old and new curriculum materials was completed for grade 2, 5, and 7 mathematics comparing: (i) the total number of topics and the corresponding time allocation, (ii) the number of subtopics, and (iii) the number of learning outcomes and learning objectives. Second, a survey was conducted with over half of the teachers that participated in the pilot to understand teachers' perception in relation to (i) the hours spent in each topic in the curriculum, (ii) the extent to which the topic is covered, (iii) the items covered in the classroom, and (iv) how well the learning purpose was achieved. The survey covered all teachers and was conducted in May 2022, towards the end of the first school year of implementation - gathering inputs from 260 mathematics and 162 science teachers. Third, nineteen focus groups were carried out by the international experts with grades 2, 5, 7, and 10 teachers (approximately 5 to 8 teachers each) who piloted the new curriculum during the 2021-2022 academic year to discuss curriculum overload among other topics.[66]

The comparison of old and new curricular documents shows that content has not been significantly reduced in terms of either the number of topics covered or expected instructional hours allocated by topic. For example, for grade 2 mathematics, the number of topics has increased from 4 to 6 while the total number of items has been reduced from 48 to 34. For grade 7 geometry, the number of topics increased from 3 to 4 while the number of items slightly decreased from 31 to 30. For grade 7 algebra, however, the number of topics increased from 6 to 7 and the number of items

65 OECD, 2020.

66 Focus groups were organized by subject and grade. There were the following: (i) Grade 2 Mathematics, Me & the Surrounding World, and ICT, (ii) Grade 5 Mathematics, Science, and ICT, (iii) Grade 7 Mathematics, Physics, Chemistry, Geography, Biology, and ICT, (iv) Grade 10 Mathematics, Physics, Chemistry, Geography, Biology, Science, and ICT.

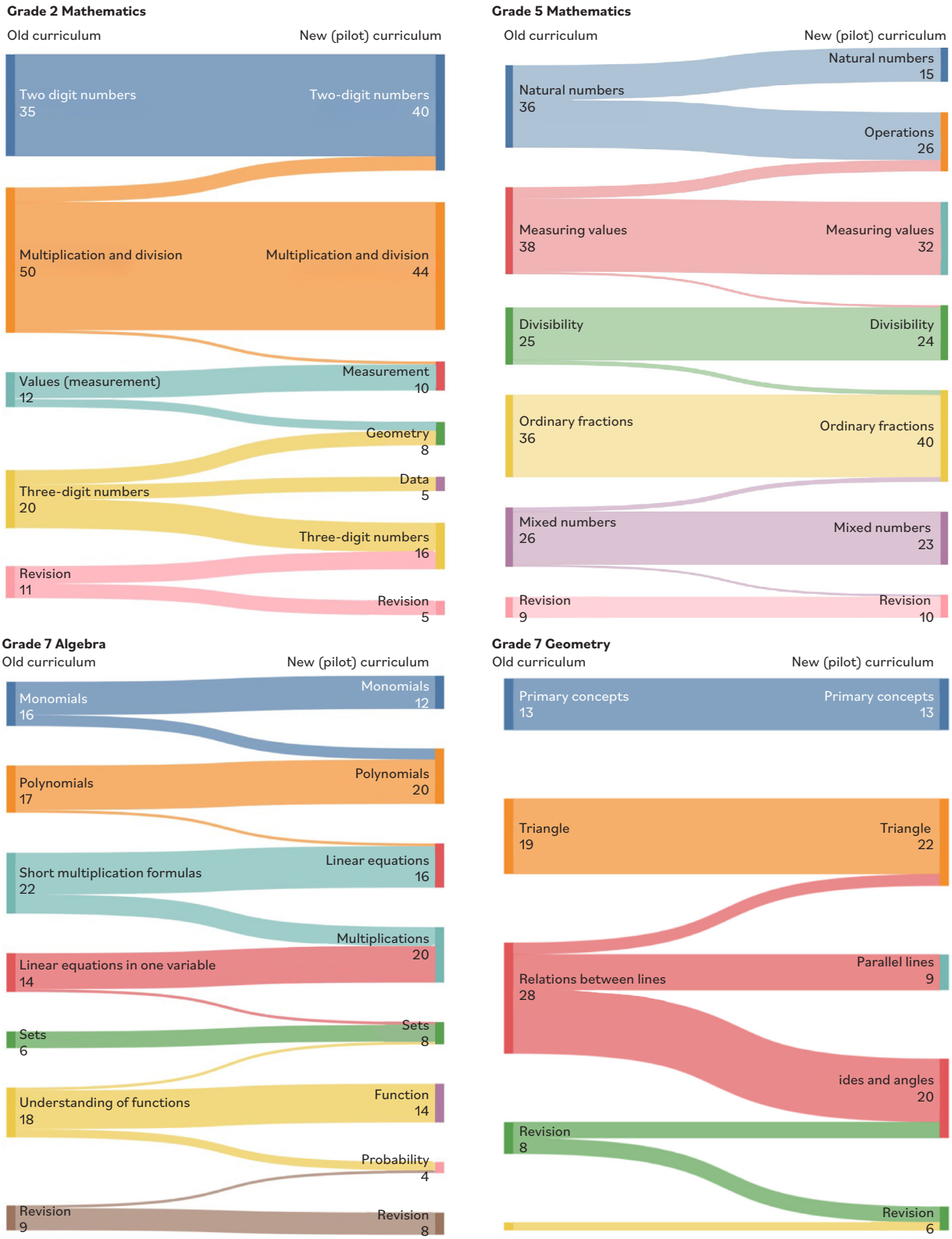
Figure 16. Content covered for each of the topics has decreased on average, but new topics were added



Source: Authors' review and analysis of old and new (pilot) curriculum documents.
Note: Topic names have been condensed for simplicity in presentation.

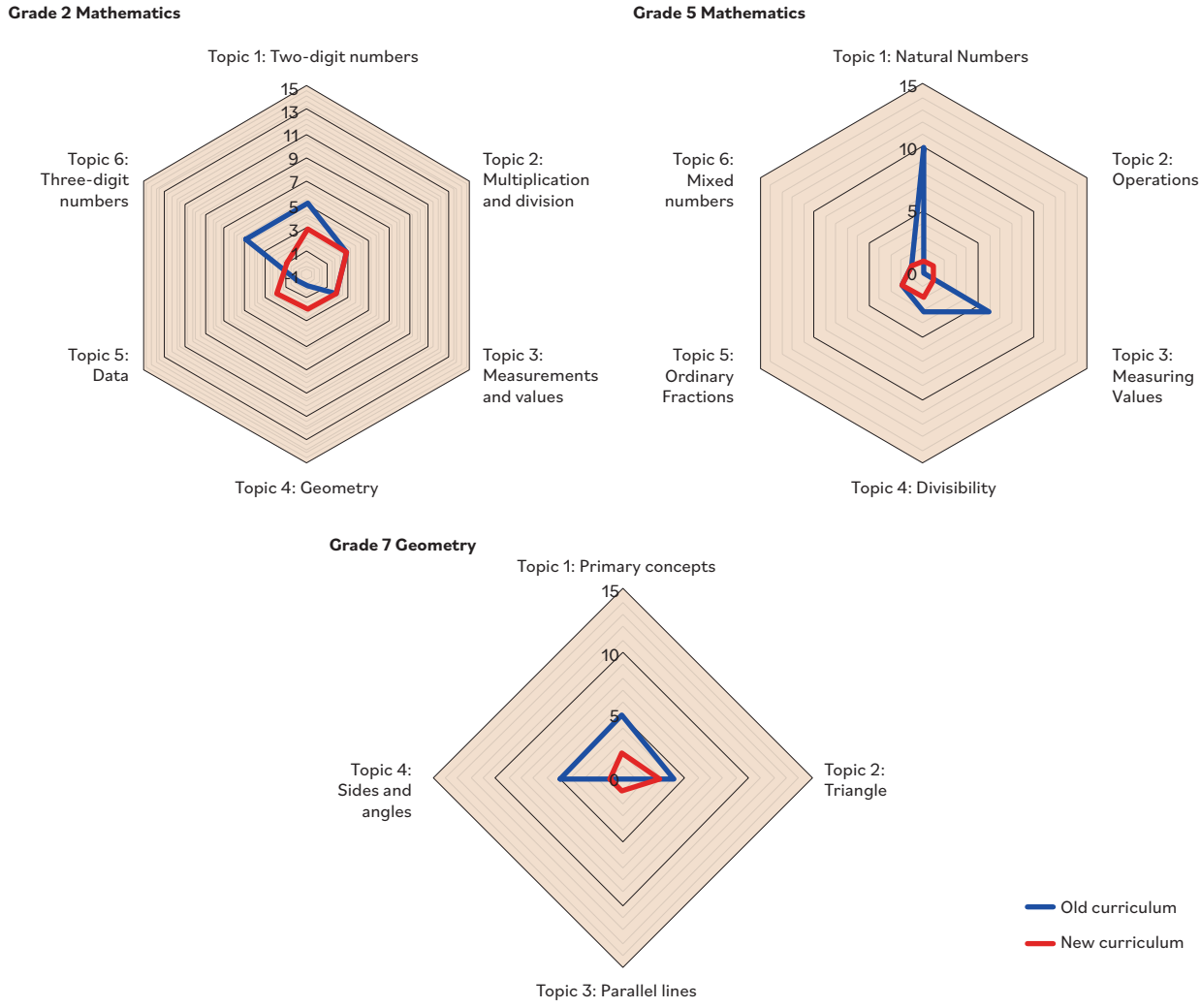
increased substantially from 26 to 37 (Figure 16). In addition, there have been changes to the time allocation to each topic. For example, for grade 5 mathematics, some topics – such as Divisibility of Numbers which went from 11 to 6 sub-topics but only reduced its time allocation by one hour – may be treated in more depth in the new curriculum (Figure 17). Other topics, such as Fractions which went from 8 to 11 sub-topics while only gaining four hours in time

Figure 17. Differences between the old and new curriculum in the allocation of hours across topics were modest



Source: Authors' review and analysis of old and new (pilot) curriculum documents.
Note: Topic names have been condensed for simplicity in presentation.

Figure 18. Learning objectives seems to have remained the same or decreased



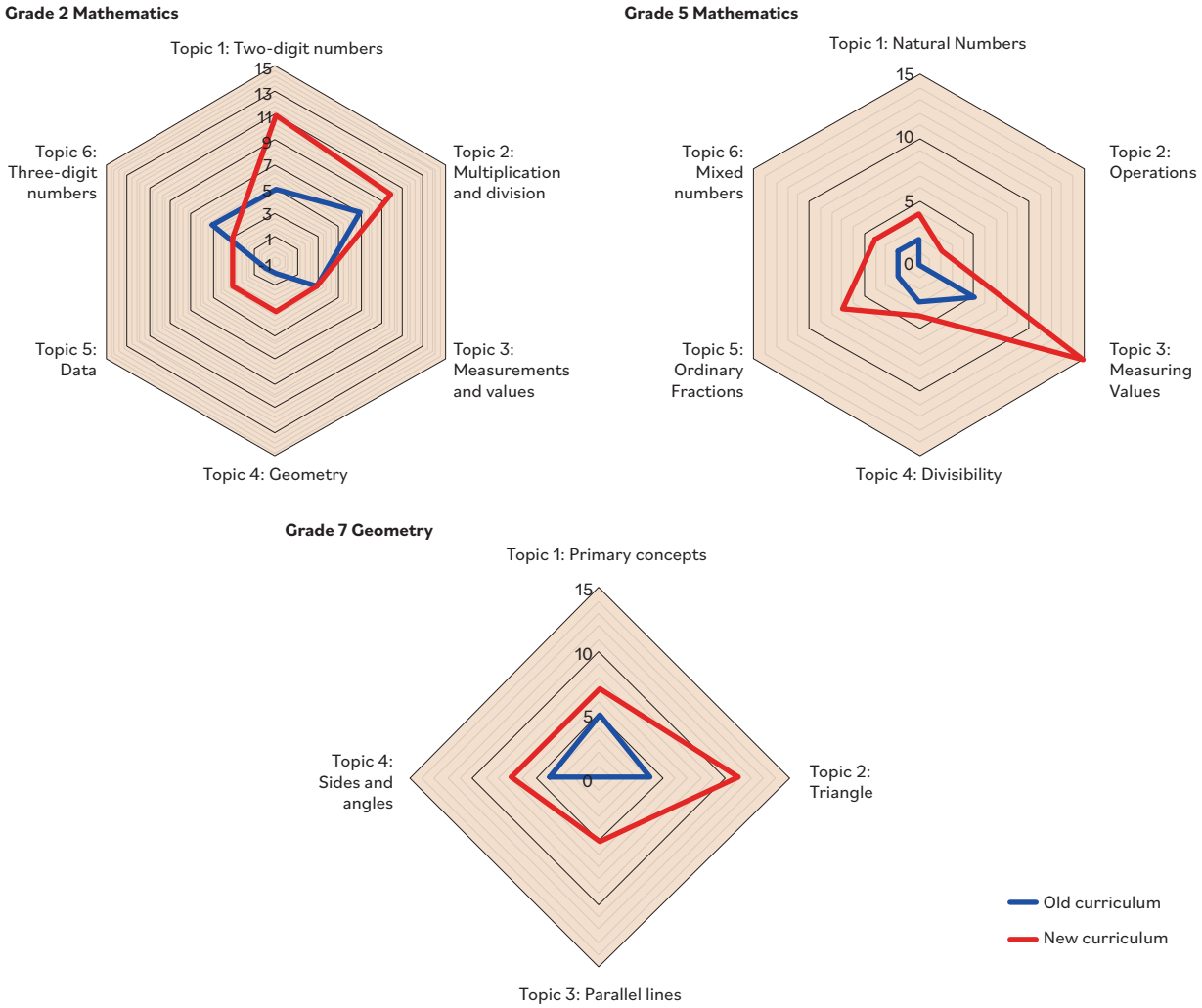
Source: Authors' review and analysis of old and new (pilot) curriculum documents.
Note: Topic names have been condensed for simplicity in presentation. Grade 7 Algebra in the old curriculum does not include learning objectives so comparison not possible.

allocation, may be treated in less depth in the pilot curriculum.[67] It is also observed that main teaching and learning objectives have decreased for all topics while learning outcomes have on averaged increased due to the sharp increase in the topic of measuring value.

The same comparison of old and new curricular documents shows that learning objectives have generally remained the same or decreased, but learning outcomes have substantially increased in most subjects, suggesting the new curriculum was overloaded. For example, for grade 2 mathematics, the old curriculum contained 15 learning objectives,

67 However, it is important to note the limitations of this type of comparison. Although current and pilot documents are similar in their contents, the level of information, detail, and sequencing can vary, leading to noise in a side-by-side analysis of both documents. In addition, the comparison of curriculum items considers the breadth but not the depth of each topic. In addition, a comparison of the total number of items for each of the curriculum topics for Grade 5 mathematics (figure 16), evidenced that the actual content covered has only marginally decreased, but this is not consistent for all topics.

Figure 19. **Learning outcomes increased on average across most grades and subjects**

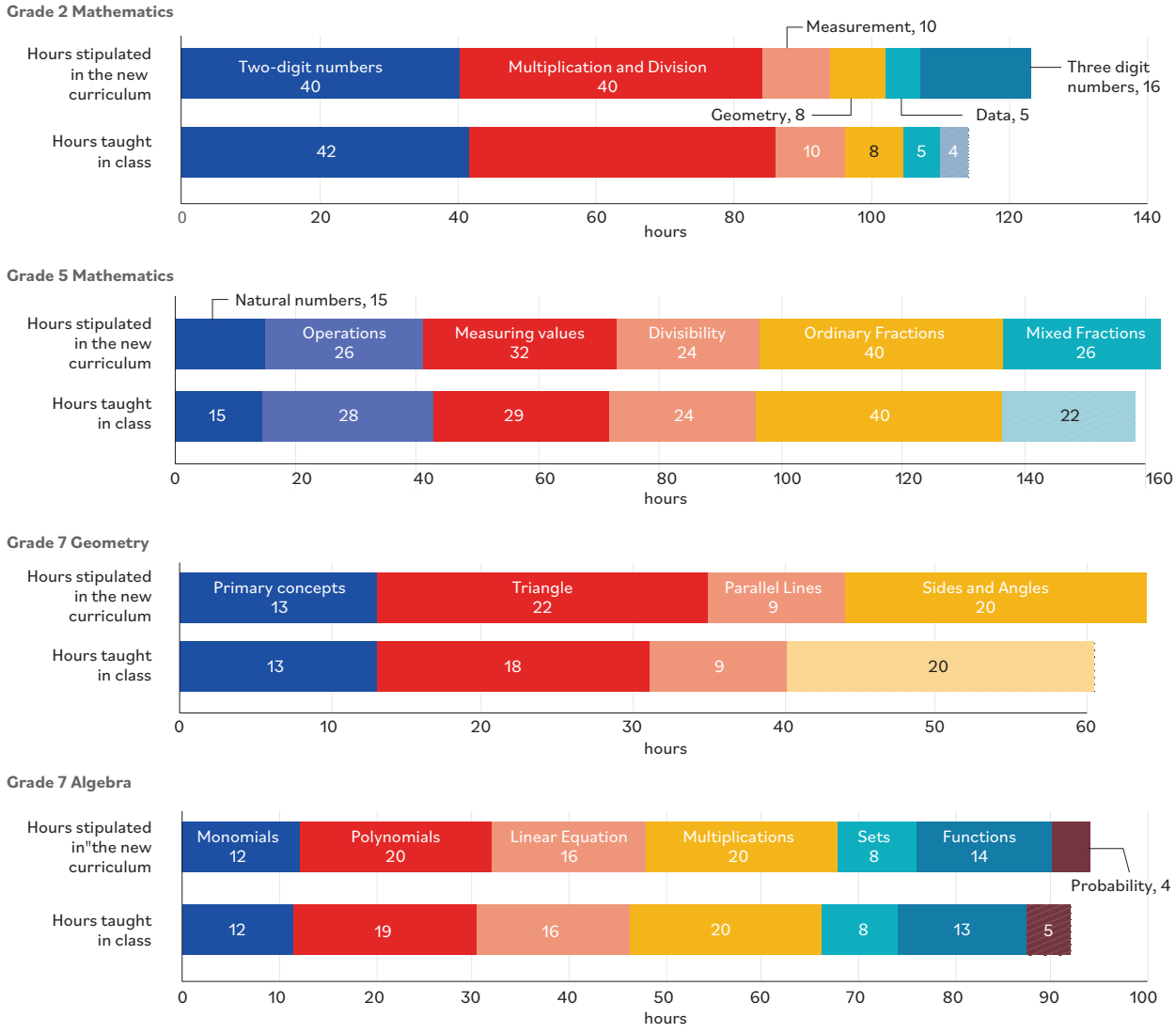


Source: Authors' review and analysis of old and new (pilot) curriculum documents. Topic names have been condensed for simplicity in presentation. Grade 7 Algebra in the old curriculum does not include learning objectives so comparison not possible.

while the old curriculum contains 13 learning objectives. A similar pattern is observed in grade 5 mathematics and grade 7 geometry, where the number of learning objectives in the old curriculum were 22 and 14 and the number of learning outcomes in the new curriculum is 8 and 7, respectively (Figure 18). On the other hand, learning outcomes were 20, 14, and 13 for grade 2 mathematics, grade 5 mathematics, and grade 7 geometry in the old curriculum, and these numbers increased substantially to 33, 36, and 30 in the new curriculum, respectively (Figure 19). Taken together this analysis suggests that the initial redesigned curriculum was overloaded.

The results of a survey with teachers suggest that while it seems they are keeping pace with the new curriculum, their confidence in the adequacy of the time allocated to each topic declined over time. In terms of time spent per topic, teachers reported being able to keep pace with the hours stipulated in the pilot curriculum across grade 2 and 5 mathematics and grade 7 geometry and algebra. For example, grade 5 mathematics teachers report to have on average spent the number of hours teacher each topic as the number of hours stipulated in the curriculum, with exception of one topic related to Operations in which an additional 2 hours was spent over the 26 hours stipulated in the curriculum.

Figure 20. Teachers report keeping pace with the new curriculum in terms of the number of hours devoted to each topic



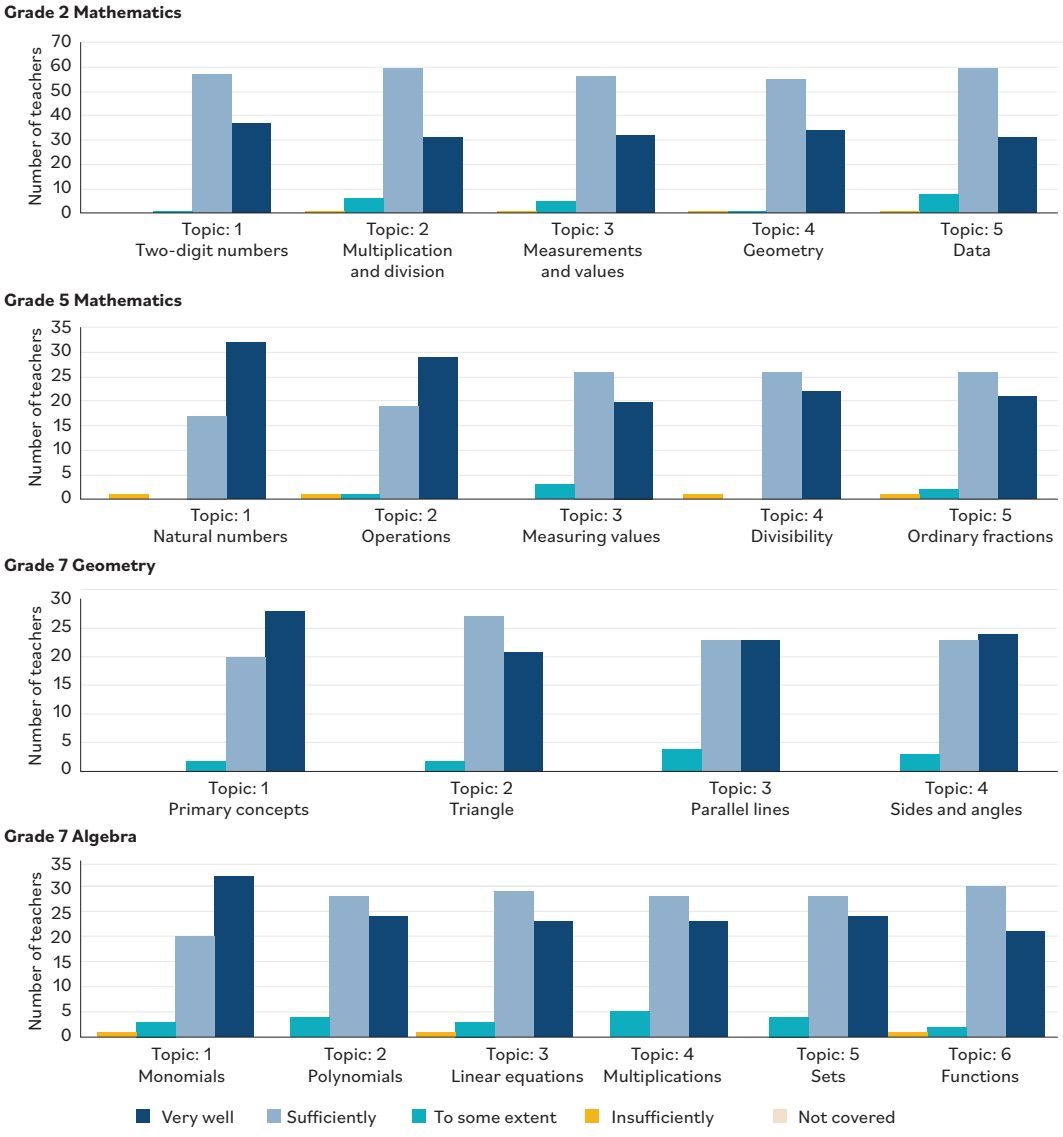
Source: Authors' calculations using Teachers Survey Data.

Note: Topic names have been condensed for simplicity in presentation. The last topic for all grades was being taught at the time the survey was conducted, thus the data is incomplete.

However, as teachers made their way through the new curriculum, their confidence in the adequacy of the time allocated to each topic declined. The number of teachers that considered the topics very well covered decreased by over a third between the first and the last topic in the curriculum (Figure 21). Teachers also expressed declining confidence that learning objectives were being met as the curriculum progressed.

The focus group sessions conducted with teachers who implemented the pilot program provided valuable insights into the potential issue of overload. During these sessions, teachers in certain subject areas expressed concerns about the curriculum being too extensive. For instance, teachers reported that the Geography curriculum contained an excessive amount of content, resulting in students memorizing the material only to quickly forget it. Similarly, many grade 10 Biology students were unable to achieve all the learning objectives due to the overwhelming amount of content. Furthermore, teachers mentioned that they had to skip or reduce practical work to prioritize the theoretical

Figure 21. **Teachers' assessment of how well covered topics are declines over the course of the school year**



Source: Author's calculations using Teachers Survey Data.

Note: Topic names have been condensed for simplicity in presentation. The last topics for Grades 2 Mathematics, Grade 5 Mathematics, and Grade 7 Algebra were being taught at the time the survey was conducted, thus the data was not collected.

aspects of the curriculum, which contradicts the intended purpose of the new curriculum. In terms of time allocations, some teachers reported that their principals had significantly reduced the assigned instructional time. For example, in Chemistry and Science, the teaching hours were reduced from 6 hours per week to just 2 hours per week due to the constraints of the school schedule. Grade 7 Biology teachers expressed the need for 3 hours per week instead of 2 to effectively teach the new curriculum. Similarly, grade 7 Chemistry teachers felt that the combination of lab work, project work, and formative assessment left insufficient time for teaching and learning the core content. While many teachers suggested that the curriculum should focus on fewer topics taught in greater depth, it proved challenging for them to identify and agree on which content could be removed to lighten the curriculum. This issue was discussed not only in the focus groups but also in the subject-specific sub-committees responsible for developing the new curricula.

Based on the results of this analysis, the STEM Technical Committee provided recommendations to reduce the curriculum load. These recommendations were followed, leading to a revision of the new curriculum for all subjects and grades. The goal of the reduction was to create a curriculum with a logical progression, starting from the highest grade of a particular subject to the first grade. This approach ensures that the highest grade aligns closely with what students are expected to demonstrate in their final examinations. To achieve this, a logical approach was taken to identify which standards (i.e., topics or skills) should be removed or taught at a lower level. The subject-specific sub-committees worked closely with international consultants and NCEDI to revise the curriculum for all subjects and grades during the 2022-2023 school year.

Assessing the implementation of teacher training

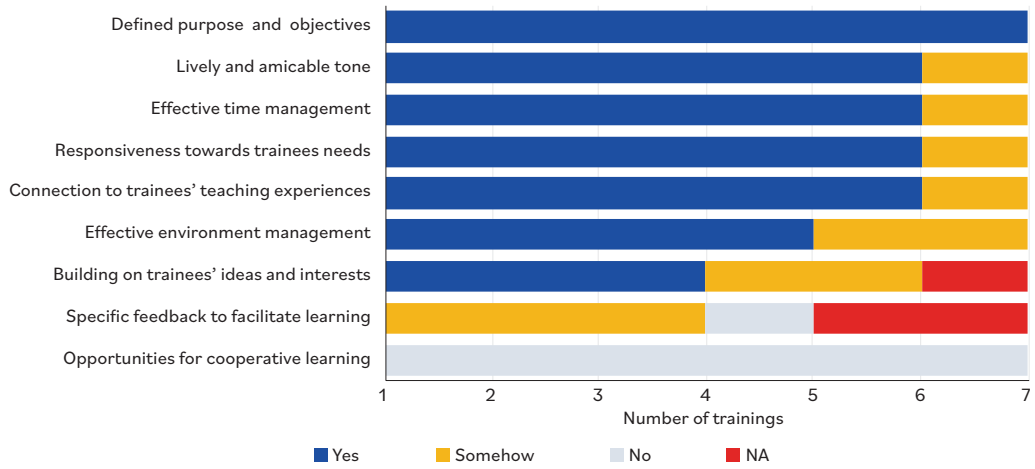
To assess the quality of training provided to teachers and school administrators, online training sessions were recorded and systematically analyzed, and feedback was gathered from participants. Online sessions of subject-specific and pedagogical training sessions were recorded and analyzed using a three-domain rubric: (i) planning and preparation, (ii) the classroom environment, and (iii) instruction. Participant feedback was collected immediately following in-person training sessions.

Analysis of recorded online sessions shows that trainers were well prepared and utilized effective instructional practices, created a good classroom environment, and adequately introduced the new standards, but the online format created challenges for the learning environment. Seven sessions were analyzed. Overall, sessions seemed to be well planned with defined objectives and effective time management, and trainers consistently demonstrated good knowledge of the content and pedagogy of the reform, as well as their subjects more broadly (Figure 22). Trainers were also well-versed in training resources and developed specific presentation material for these workshops. The trainers effectively managed time, responded to trainees' needs, and created an approachable and dynamic classroom environment. They used real-life teaching examples to clarify content and answered questions in detail while maintaining the pace of the agenda. In some cases, sessions were split into two parts to allow for equal time for presentation and discussion. All observed training courses introduced the new standards for the subject and referred to the overall principles of the reform. The trainers highlighted connections between new learning standards and made direct comparisons between the previous and new standards. Taken together, the results show the online training sessions were delivered well, but that there were missed opportunities for cooperative learning as most training sessions were delivered as lectures without much engagement beyond opening for questions or comments. Live interaction mostly occurred at the end of the presentation or during the dedicated time for questions and answers likely due to the online format of the training.

Teacher and school principal opinions collected immediately after in-person teacher training, both pedagogical and subject-specific, as well as administrative staff training, were consistently positive.[68] These surveys were designed with yes/no categorical responses to gauge overall satisfaction with the training immediately after its completion. First, in terms of the in-person pedagogical teacher training, all the trainers received a ranking above 9 on a scale of 1-10. Regarding the content, 95 percent of participants reported that the training session contributed to improving their knowledge. Additionally, 93 percent felt that the tasks were appropriately timed, and 98 percent believed that the materials presented were in line with the topics advertised in the agenda. Secondly, for the in-person subject-specific teacher training, 97 percent of participants reported that the training session contributed to improving their knowledge. Furthermore, 95 percent considered the training to be relevant, and 88 percent found it easy to understand. Lastly, for the administrative staff training, 89 percent of respondents expressed that the workshop contributed to their knowledge. Additionally, 78 percent found the content comprehensible, 78 percent considered it relevant, and 80 percent felt that

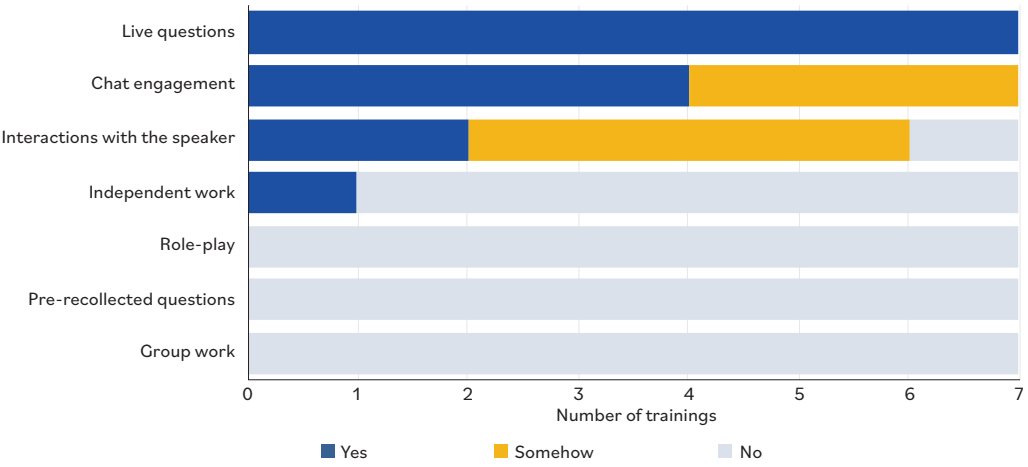
68 AYB Educational Foundation (2022).

Figure 22. There were a number of positive features of training delivery were present across most of the 7 training sessions analysed



Source: Authors' coding and analysis of recorded training sessions.

Figure 23. Opportunities for engagement during online training sessions were limited



Source: Authors' coding and analysis of recorded training sessions.

there was sufficient time for questions. Overall, the feedback received from participants in all three types of training was overwhelmingly positive, indicating a high level of satisfaction with the training sessions.

During focus groups, teachers also reported that the mentoring scheme had a positive impact on their experience and self-confidence and was essential for their success in delivering the new curriculum. Tavush teachers involved in the pilot were also able to express their feedback in 19 dedicated focus groups that discussed teacher training and mentoring among other topics. Teachers reported that the weekly mentoring sessions to review lesson plans and assessment methods were greatly beneficial. During these sessions, teachers received helpful instructional materials from their mentors that successfully dissipated their lack of confidence with topics (e.g., new topics in the curriculum such as financial literacy in grade 5 math). This mentoring also improved their understanding and implementation of formative assessment.

Assessing teachers' and principals' behaviors, practices, and perceptions about the new curriculum

To understand a range of factors related to behaviors, practices, and perceptions that might support or hinder the adoption of the new curriculum, two rounds of surveys were carried out at the end of the 2021-2022 and 2022-2023 school years with teachers and school principals. Teachers are the direct implementers of curriculum reforms, and as such, their perception of the process and buy-in on the reforms shape their actual implementation.[69] Evidence suggests teacher's buy-in of a reform during its early stages can be used as proxies to measure successful future implementation and effectiveness.[70] School principals can also play an important role in facilitating and managing the adoption of new pedagogical practices at the school. As such, surveys were applied to school principals and STEM teachers in pilot grades after the initial pilot was implemented in Tavush.

Indices of satisfaction with the teaching profession, wellbeing, classroom practices, management practices, perceptions about the new curriculum, satisfaction with teaching and learning materials and training and perceived obstacles to implement the new curriculum were constructed (Figure 24). The indices are generated in three steps. First, a range of questions were used to build each index and Likert-scale responses to each question were normalized on a 0 to 1 scale. For all questions, the response associated with worst practice or negative perception is normalized to 0, and the one associated with best practice or positive perception is normalized to 1. For example, when asking whether teachers and principals agreed with the following statement "The new curriculum standards will enable me to spend more time teaching higher-level (i.e., critical and creative) thinking skills", the response "Agree a lot" is ranked 1 and the response "Disagree a lot" is ranked 0. The "in between" category "Disagree a little" is assigned a value of 0.33 and "Agree a little" is assigned a value of 0.66. Similarly, for three categories, the "in between" category is assigned a 0.5. For perceptions of obstacles to implement the new curriculum, the construction of the index follows the same logic, but it is reversed so 1 is associated with no obstacles and 0 with several obstacles. Second, the indices and sub-indices for each year are calculated as the unweighted average of the normalized responses. Third, different samples of teachers were surveyed in the first and second year of implementation, and thus, comparisons between the two years do not necessarily reflect improvements or declines in each index and are not reported. For this reason, the final index for each construct is an unweighted average of the corresponding yearly index.

Figure 24. Indices on teacher's and principal's behaviors, practices and perceptions

69 OECD, 2020.
70 Hübner, Savage, Gräsel and Wacker, 2021.

Behaviors

In Tavush, on average, teacher's satisfaction with the profession ranks 0.73, where 0 represents low satisfaction and 1 represents high satisfaction, suggesting there is room for improving the overall satisfaction of teachers with both the teaching career and their current posts. Questions to capture teachers' satisfaction with their profession include whether they perceive more advantages than disadvantages of being a teacher and if they either regret or would choose again the teaching profession. Questions to capture their satisfaction with their current job focus on whether they enjoy their work at the school, if they would recommend their school as a good place to work, if they are satisfied with their performance at the school and if they are satisfied with their current job.

Teacher's mental health and wellbeing in Tavush ranks at 0.63, on average, where 0 represents low wellbeing and 1 represents high wellbeing, pointing to the need to work towards teachers' wellbeing. Questions to capture teacher's mental health and wellbeing include whether teachers have many students in class, they have too much material to cover in class, have too many teaching hours, need more time to prepare for class, need more time to assist individual students, feel too much pressure from parents, report difficulties keeping up with the changes to the curriculum, and have too many administrative tasks.

The relationship between principals and teachers index ranks at 0.80, on average, suggesting good workplace relationships within schools. Questions to capture the quality of the relationship between principals and teachers were asked to school principals only and include whether they feel supported by teachers, valued by teachers at the school, have a good relationship with teachers, are treated with cordiality and respect by teachers and their decisions are respected by teachers even when in disagreement.

Practices

In Tavush, on average, teacher's self-reported classroom practices rank 0.66, where 0 represents poor classroom practices and 1 represents best classroom practices, suggesting that teachers are implementing several practices in the classroom that might enable the adequate adoption of the new curriculum, including the use of formative assessments, while some others are yet to be implemented. Questions to capture teachers' classroom practices focus on whether teachers report giving emphasis to a range of approaches and processes the curriculum for their respective grades and subjects – knowing basic facts and principles, providing explanations of what is being learned, designing, planning, and conducting investigations or projects, doing exercises and problems, and integrating the subject with other subjects. It also captures whether teacher report using in many lessons the following methods for assessing student learning – develop and administer assessments (own or national test), have individual students answer questions in front of the class, provide written feedback on student work in addition to a grade, let students judge their own progress, observe students when working on particular tasks and provide immediate feedback, collect data from classroom assignments or homework. Related to practices on assessments, additional evidence from focus group discussions have shown that most teachers relied heavily on formative assessment, although for many teachers, considerable time was needed to adjust to this new approach. Teachers found formative assessments to be a useful mechanism for assessing students. They considered it innovative and enjoyable and thought it promoted students' engagement with learning and assessment. Given this was a new approach to assessment, teachers reported that a large number of teachers, parents, and students were nervous about it at the beginning of the year, with teachers finding it difficult to implement and not feeling prepared to do so. However, implementation became smoother with help from mentors and collaboration between teachers. This points out to the process of change embedded in the curriculum reform that requires time, space and investments to ensure all relevant stakeholders are on board and ready to implement the proposed changes inside the classroom.

Principal's management practices rank at 0.69, on average, where 0 represents poor management practices and 1 represents best management practices, which suggests that principals are already adopting several management practices that can support teachers in the implementation of the new curriculum with some room for improvement.

The dimensions measured include operations management, target setting, performance monitoring, and people management.[71] Questions to capture performance monitoring and target setting include whether student assessments are used to inform parents about their child's progress, to make decisions about students' retention or promotion, to compare the school to regional/national performance, to monitor school's yearly progress, to make judgements about teacher effectiveness, to identify aspects of instruction or curriculum that could be improved or to compare school with other schools and if achievement data us posted publicly and tracked over time by an authority. Questions related to operations management and people management include how do appraisals and/or feedback to teachers are related to changes in salary, bonuses or monetary rewards, opportunities for professional development, career advancement opportunities or public recognition, changes in responsibilities, and the frequency of school principals engaging with teachers to help build a school culture of continuous improvement, asking teachers to participate in reviewing management practices, solving classroom problems with teachers, discussing the school's academic goals with teachers at faculty meetings, referring to the school's academic goals when making curricular decisions with teachers, setting aside time at faculty meetings for teachers to share ideas or information from in- service activities, conducting informal observations in classrooms on a regular basis, among others.

Perceptions

Teachers and principals report on average some constraints that might hinder their capacity to implement the new curriculum, with teachers ranking 0.72 and principals ranking 0.58 on the index, where 0 represents many obstacles and 1 represents no obstacles, suggesting that there are important investments to be made in this area to minimize the obstacles schools face in implementing the curriculum. Questions to capture obstacles to their own implementation of the new curriculum include lack of or inadequate/poorly qualified teaching staff, lack of or inadequate or poor-quality educational material, a lack of or inadequate or poor-quality physical infrastructure.

Teachers' perceptions about the new curriculum rank at 0.69 on average, where 0 represents negative perception towards the new curriculum and 1 represents positive perception, suggesting that some investment is needed to ensure teachers understand, and are comfortable and confident about delivering the new curriculum in the classroom. The dimensions measures include preparedness, impact on education, workload, and morale, for which sub-indices are also created.[72] Teacher's perception of their preparedness level ranks highest at 0.84. Questions to capture perception in terms of preparedness to implement the new curriculum include whether they feel well informed regarding what the new curriculum standards are; are sufficiently prepared through professional development to transition from teaching current standards to the new standards, and to properly prepare for the implementation of the new standards. Teacher's perception of the impact of the curriculum on education ranks at 0.75. Questions to capture perception on the impact the new curriculum will have on students' lives and education measure whether they believe belief that the curriculum change will be more effective than current standards at preparing students for their life, feel that there is a difference between the old and new curriculum, feel the new curriculum is more positive than negative step for education reform. Teacher's perception of workload ranks at 0.50. Questions to capture teacher's perception of the workload related to the implementation of the new curriculum include whether they perceive the new curriculum to be easier to understand than the old one, that the work that will be put into preparing and transitioning to the new set of curriculum standards will be worthwhile. Teacher's perception of morale ranks at 0.67. Questions on the impact of curriculum reform on teacher motivation and perception of the teaching profession include whether teachers are concerned that the new curriculum standards will restrict their creativity and the types of instructional strategies they may use, teachers

71 We follow the school management literature in the construction of this index (Bloom et al 2015), borrowing questions from the school survey in PISA 2021 as done by Leaver, Lemos, and Scur, 2020.

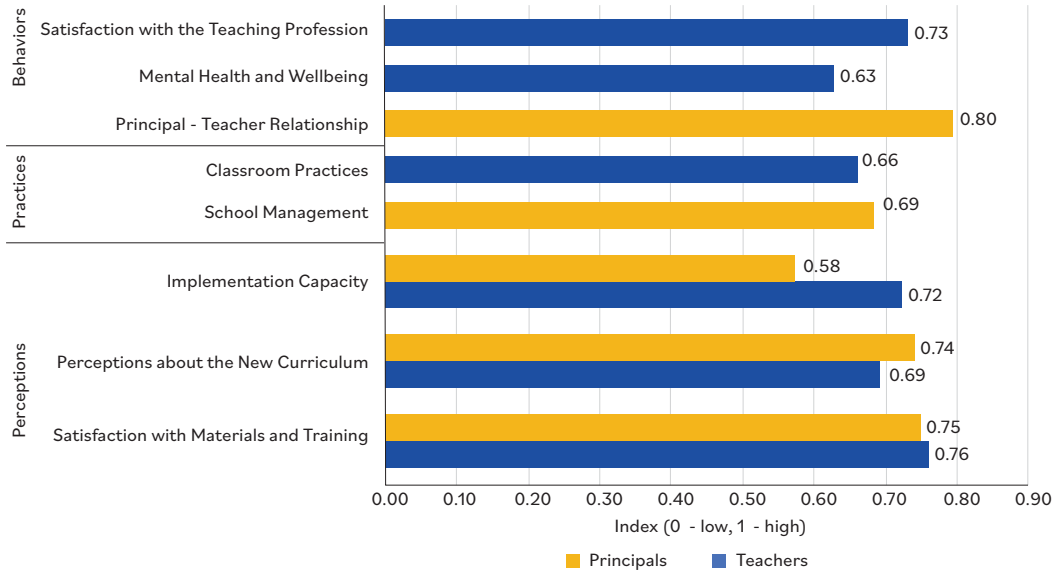
72 The surveys included a set of 16 multiple choice items to explore their perception on the curriculum reform. The items were adopted from Cheng, 2012, and Cochrane and Cuevas, 2015, who based his questions on those used by Mertler, 2011 and Smith and Kovacs, 2011 in their studies of teacher perceptions of the No Child Left Behind Act in the United States. Participants responded to each item based on a four-point Likert scale (strongly agree, agree, disagree, strongly disagree).

would encourage others to enter the teaching profession at this time, teachers would like more decision-making power over the curriculum than what they believe the new set of standards will permit, teachers perceive the new curriculum standards to help them become a more effective teacher.

Principals' perceptions about the new curriculum rank at 0.74 on average, following the same set of questions asked to teachers and suggesting an overall better perception of school principals relative to teachers. Principals' perception of their preparedness level ranks highest at 0.86. Principals' perception of the impact of the curriculum on education ranks at 0.82. Principals' perception of workload ranks at 0.56. Principals' perception of teacher morale ranks at 0.71.

Teacher's and principals' satisfaction with the new teaching and learning materials and training rank at 0.76 and 0.75, respectively, where 0 represents negative perception and 1 represents positive perception, suggesting an overall positive perceived quality of teaching and learning materials and training received despite some issues reported during the first year of the pilot. Questions to capture satisfaction with the teaching and learning materials measure whether teaching and learning materials are in line with the curriculum and whether teachers are satisfied with the teaching and learning materials provided, and questions about satisfaction with laboratories. The index also includes satisfaction with training offer, including if the training offer for implementing the new curriculum is enough to properly implement it, and if the offered trainings are a good preparation for implementing the curriculum. The results show a positive perception of the teaching and learning materials and trainings but focus groups discussions showed there were some challenges with the timely delivery and utilization of printed teaching and learning materials, mainly due to the complexities of producing them, the relatively short timeline for the pilot, and financing issues. Teachers reported that most teaching and learning materials including textbooks, computers, and laboratory equipment were not ready for use in the first year of the pilot. For example, teachers pointed out that the materials initially developed for digital literacy and computer science had to be adjusted as they were exceedingly difficult. Since most schools did not receive textbooks on time, teachers relied instead on other materials which included old textbooks, printouts, and handouts of activities for in-class instruction. These challenges, paired with internet connectivity problems, contributed to the delays in the timely delivery and utilization of printed teaching and learning materials and the use of online teaching

Figure 25. Teachers' and Principals' indices to measure behaviours, practices, and perceptions about the curriculum reform



Source: Authors' estimations using survey data from teachers and principals in Tavush in 2021-2022 and 2022-2023 school years.

and learning materials. On the other hand, focus groups with teachers highlighted the mentoring designed to support teachers throughout the year, and teachers reported that it had a positive impact on their experience and self-confidence.

The results of this survey highlight areas where additional investments are needed and can guide nationwide implementation moving forward. Taken together, the evidence shows that teachers are moderately satisfied with their jobs and principals feel content with their relationships with teachers, showing a nurturing environment at schools. They seem to have adopted on average adequate classroom and management practices that can facilitate or enable the implementation of the new curriculum, albeit with some room for improvements to strengthen these practices. However, teachers and principals feel that there are some obstacles related to the shortage and adequacy of teaching staff, education materials, and school infrastructure that may hinder their capacity to implement the new curriculum. Finally, teachers and principals view the curriculum reform positively, and feel they are mostly prepared for its implementation, and that the reform can have an important impact on the performance of the education system. However, considerations must be made to ensure teachers do not feel this to be an additional burden to their workload and ensure teacher morale and mental health and wellbeing are high (Figure 25).

7. Evaluating the impact of the new curriculum on learning

Establishing a robust evaluation strategy

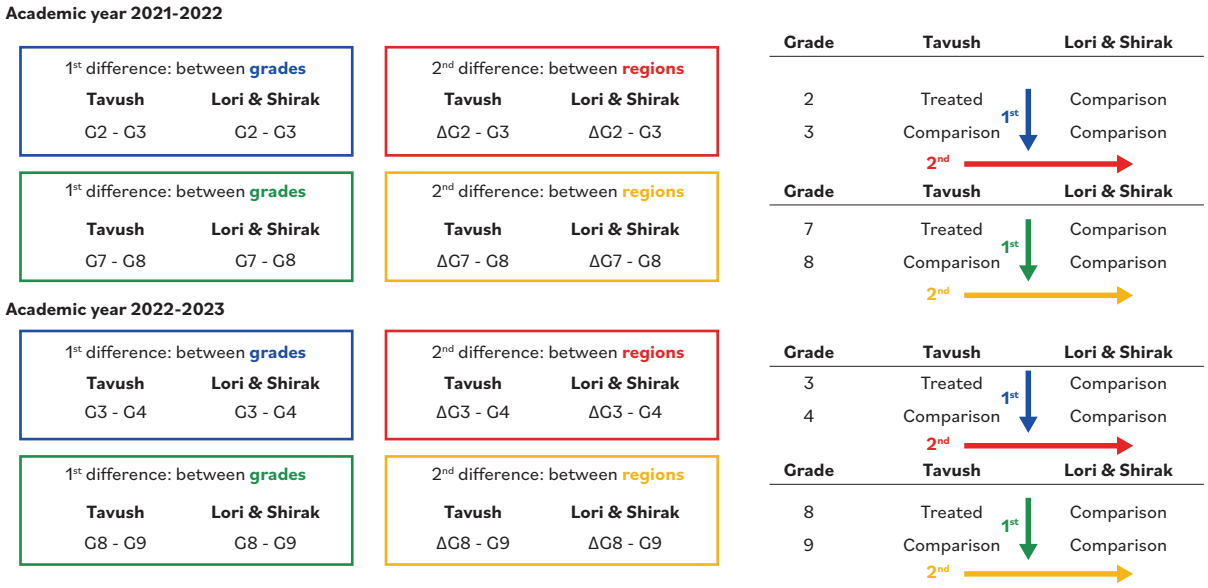
To estimate the impact of the reform, we exploit variation in the regions affected by the policy and in the grades targeted by the curricular reform following a differences-in-differences approach. The first difference is the difference in outcomes between consecutive pairs of grades, where one grade was not targeted by the reform (for example, grades 4 and 9 in the academic year 2022-2023) and the other was (for example, grades 3 and 8 in the academic year 2022-2023). The second difference is the difference in the outcome gap between pairs of grades across the treatment and comparison regions (Figure 26).[73]

This method relies on the assumption that the knowledge gap between cohorts in consecutive grades would have stayed the same in the absence of the curriculum reform. There is supporting evidence in favor of this assumption, known as parallel trends, for the region groups and subjects selected as part of the evaluation strategy. First, Lori and Shirak were selected as an ideal comparison group since they were the regions that most resembled Tavush in trends over the last decade, size, and school characteristics. Both groups have around 14 classrooms and 215-230 students per school, similar school services, and comparable proportions of Basic and Secondary schools (Figure 27). Second, historical data suggests similar levels and trends across Tavush and the combination of Lori and Shirak in regional mathematics and physics assessment results. Event-study coefficients and 95 percent confidence interval intervals[74] capture that the difference in trends across regions is not statistically different from zero (Figure 28). It is important to note that when comparing to Yerevan, the capital has more students and schools, and outperforms the other regions in the assessments.

73 The differences-in-differences estimations includes a range of controls such as variables for student gender, school infrastructure, rural location, the logarithm of the number of students, the logarithm of the number of teachers, and the date of the assessment. Additionally, standard errors are clustered at the school level.

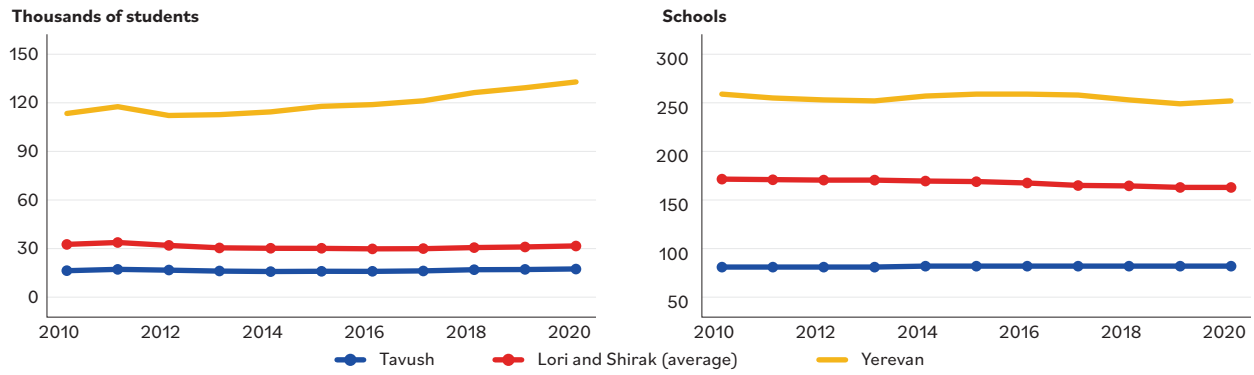
74 95 percent confidence intervals are computed since the data comes from a sample of students, that is, if one were to assess the total student population instead of a sample, the true parameter would lie in this range with 95 percent certainty.

Figure 26. A difference-in-differences evaluation design was followed to evaluate the impact of the curriculum reform



Source: Authors' presentation. Grades 2 and 7 were treated in the academic year 2021-2022 and Grades 3 and 8 were treated in the academic year 2022-2023. The consecutive grades were not treated in those respective years.

Figure 27. Historical trends in school characteristics are similar across Tavush and comparison regions

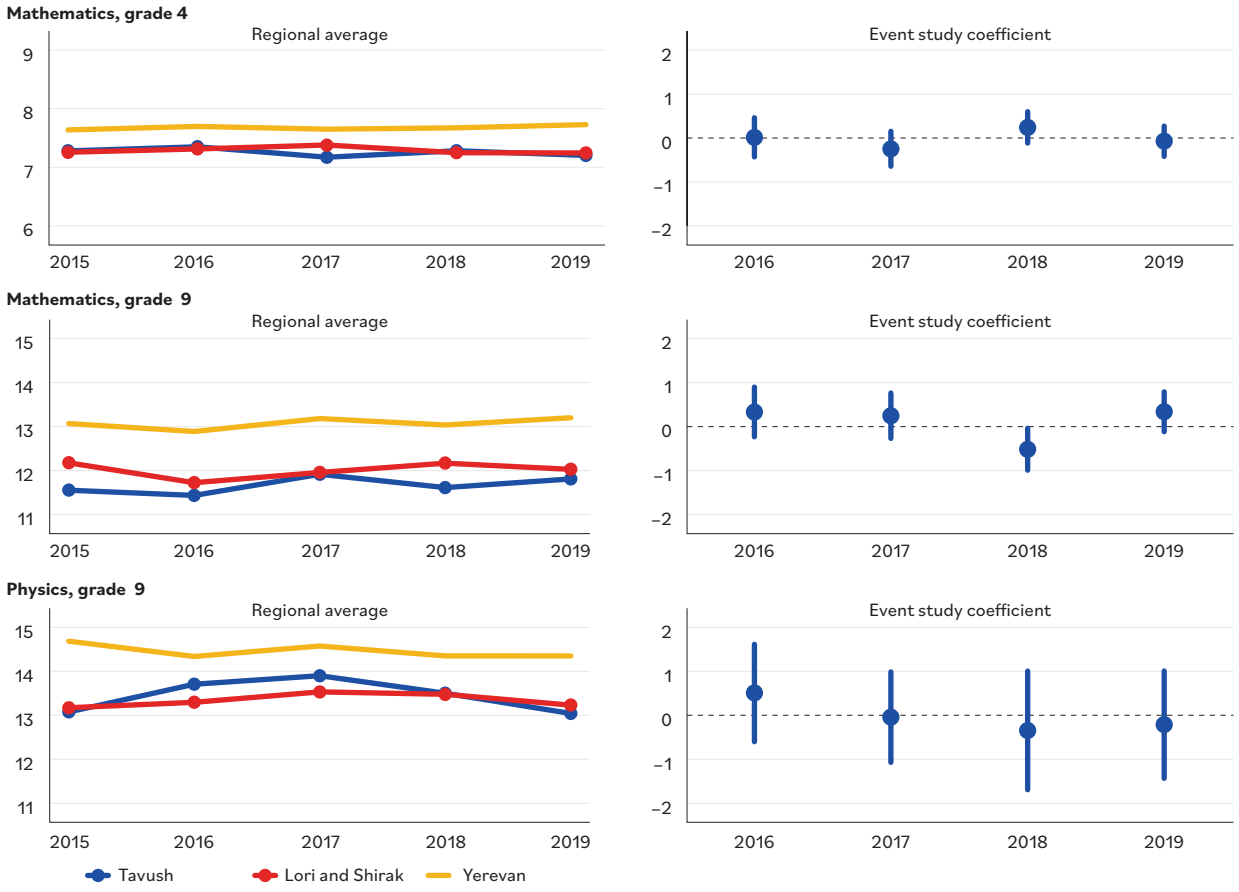


Source: Authors' estimation using data provided by NACET Statistical Database. Yerevan is used for benchmarking only.

The evaluation is focused on schools that offer education in grades 1-12 in Tavush, Lori, and Shirak, and on students that were in grades 2, 3, 7, and 8 in the 2021-2022 school year and 3, 4, 8 and 9 in the 2022-2023 school year. In the first round, there were 144 schools and 6,756 students who took the mathematics or science assessment. In Round 2, the evaluation included 207 schools and 9,981 students.

The average student is 9.3 years old for 3rd and 4th grade, and 13.2 for 8th and 9th grade. Overall, 46 percent of students are female, 89 percent have indicated to mostly speak Armenian at home, and 19 percent have parents who completed secondary education. In terms of household goods, 74 percent have a computer or tablet and 91 percent have an Internet connection at home (Figure 29). Most teachers are female, 45 percent are between 25

Figure 28. Historical trends in end-of-year school-level students evaluations are similar across Tavush and comparison regions



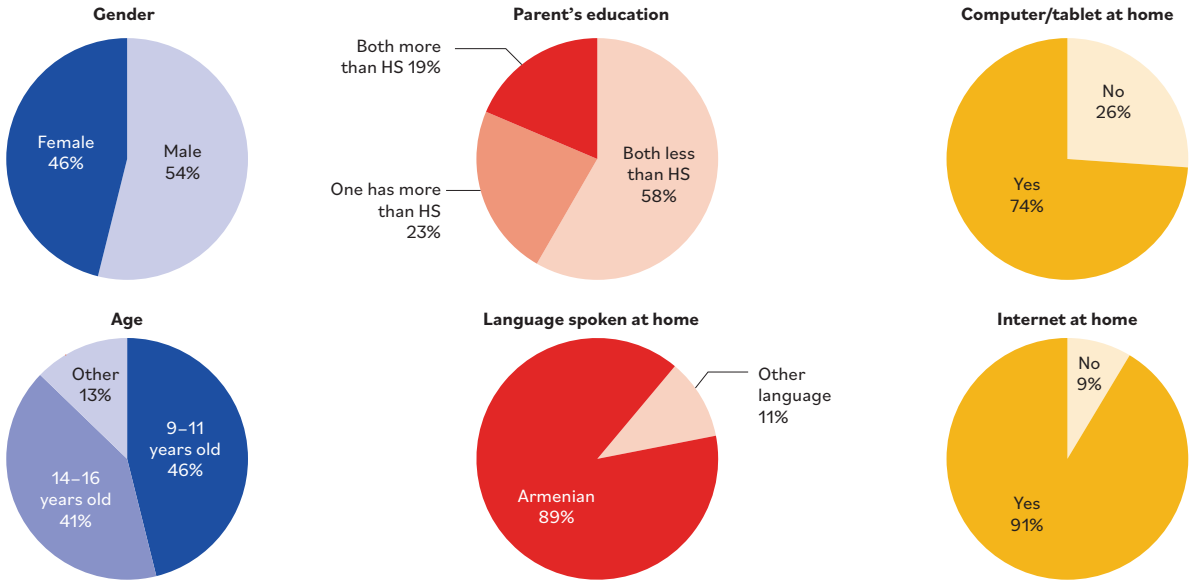
Source: Authors' estimation using data provided by Assessment and Testing Center Statistical Database. The confidence intervals in the figures on the right are represented by the vertical lines and include zero in all cases. Yerevan is used for benchmarking only.

and 50 years old, 54 percent have a master's degree and 62 percent have teaching experience in one school (Figure 30). As for the principals, the majority are female, over 50 years old, hold a master's degree, and have experience working as a principal in one school (Figure 31).

What is the overall impact of the curriculum reform?

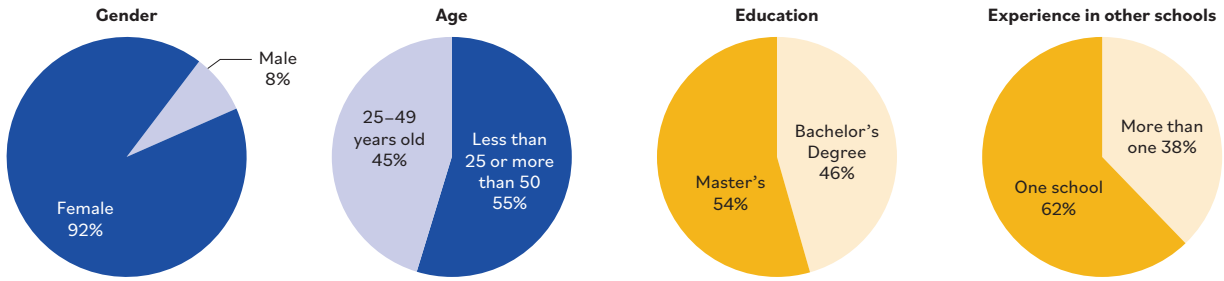
The impact of the curriculum reform is assessed using student assessment data collected for mathematics and selected science subjects, as described earlier. Both *item response theory* (IRT) and *classical test theory* (CTT) scores are used for an initial assessment, but most results presented focus on the IRT scores, which are standardized relative to the combined distribution of the adjacent grade pairs. Box 3 explains the use of IRT scores. For completeness, effects of the reform on CTT scores are also shown, which are constructed as the share of correct answers and standardized relative to the comparison group (non-target grades in Tavush, and all of Lori and Shirak). While IRT scores are preferable for the reasons outlined below, CTT scores are often used by assessment agencies. By standardizing the scores, effect sizes across subject areas and years can be compared.

Figure 29. The average student in Tavush and comparison regions is male, has parents with secondary education and has access to tech devices and internet at home



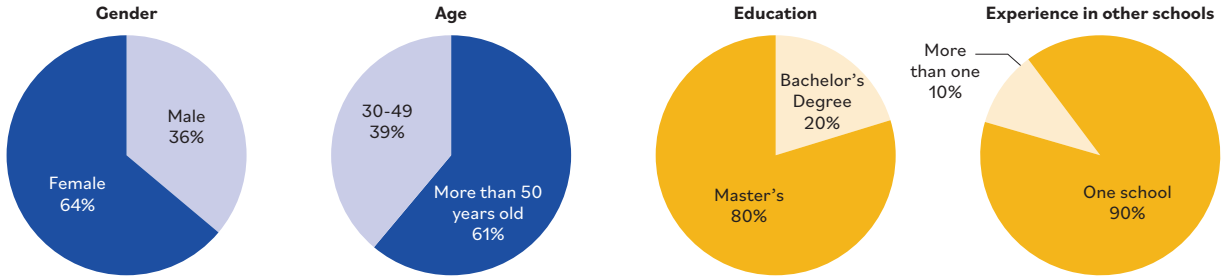
Notes: Information from Round 2 of the data collection. Full sample of students.

Figure 30. The average teacher in Tavush and comparison regions is female, has a master's degree, and work experience in one school



Notes: Information from Round 2 of the data collection. Sample of teachers that could be matched with their students.

Figure 31. The average principal in Tavush and comparison regions is female, has a master's degree, and work experience in one school



Notes: Information from Round 2 of the data collection. Sample of principals.

Box 3. The use of Item Response Theory for computing scores of student assessments.

Item Response Theory (IRT) was used to calibrate items and to generate comparable scores. Because of the quality of the algorithms implemented in the program, and the familiarity of its use, it was decided to use PARSCALE (version 4.1; Scientific Software International, 2003) for the IRT analyses, and the Graded Response Model was chosen, given its flexibility and quality of the output. A two-parameter logistic (2PL) model was used (difficulty and discrimination parameters). This model is suggested for the national exams because it provides all necessary information for test development regarding item difficulty and item discrimination, as well as information of the item, and the test information function. Specifically, the method of scoring subjects was "expectation a-posteriori" (EAP; Bayes estimates). Type of prior: Normal approximation. The Score Mean= 50 (SD=10.00). Number of quadrature points: 30.

The item difficulty parameter describes the inflection point of the item function logistic curve. It coincides along the ability scale with the point where the item has a probability of 50 percent correct responses, when the guessing parameter is equal to 0.

Item discrimination is given by the slope of the ICC at the inflection point: the steeper the curve, the better the discrimination of the item. Item discrimination indicates how well it differentiates examinees with ability (or measured trait) below the item location from those with ability above the item location.

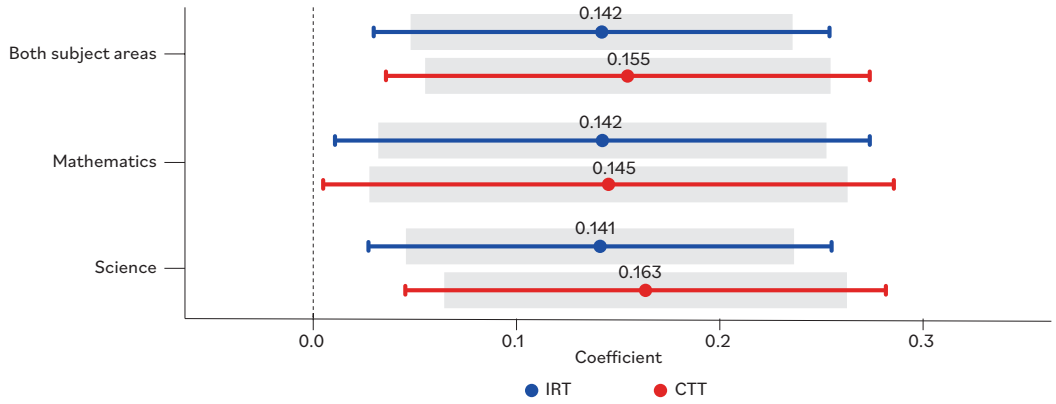
Scaling and linking: Two equating methods were applied in order to 1) analyze the quality of the items (estimation of parameters) and obtain comparable individual students' scores, and 2) compare the general performance between different consecutive grades of the same year, and consecutive grades across the two data collection rounds (2022 & 2023).

Firstly, vertical scaling was carried out for each Grade considering the first grade as the reference group (e.g. Grade 2 for Grade 3 in Math exam). Therefore, as both groups are non-equivalent samples, but they answered common items (the overall exam or some anchor items), all parameters of the first group were used for running the calibration of items of the second group, in order to put this last group in the same scale as reference group. The resulting scores of this equating procedure are more precise and useful to interpret at individual level.

Secondly, concurrent calibration was used with a non-equivalent groups anchor-test (NEAT) design for equating, in order to calibrate all items to the same scale. In this approach, all the item parameters for the items in two or more groups (e.g. two consecutive grades[75] of Year 2022; the same group of students in 2022 and 2023), are estimated simultaneously in a single calibration run. It allows to capitalize on common items across the two data collection rounds (within subjects and grades) to map assessment results onto the same scale. Because the two exams for both groups have items in common, the resulting item parameters for all items included in the concurrent calibration run are on the same scale (Kang & Petersen, 2012). In this way, results from all of these different administrations can be compared as they are all expressed on the same scale. Therefore, we obtained a single group distribution (with different parameters than the previous equating) resulting in comparable scores to be interpretable at group level. IRT-scaled scores were standardized to have a mean of 0 and a standard deviation of 1 with respect to each grade pair and year.

75 Hambleton, Swaminathan and Rogers, 1991, p. 15

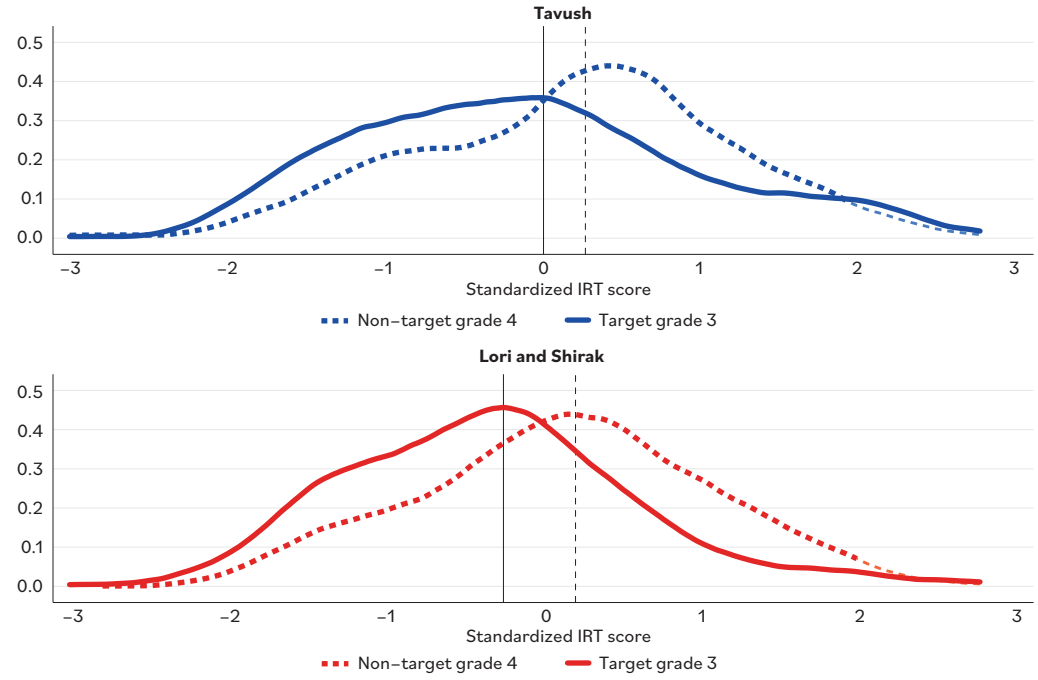
Figure 32. The impact of the curriculum reform was positive, overall and by subject



Source: Authors' estimations.

Notes: All students, appended by rounds of data collection. Because the data for the impact evaluation comes from a sample of students, we compute confidence intervals. That is, if one were to assess the total student population instead of a sample, the true parameter would lie in the range delimited by the red and blue lines with 95 percent certainty, and in the range delimited by the gray bar with 90 percent certainty.

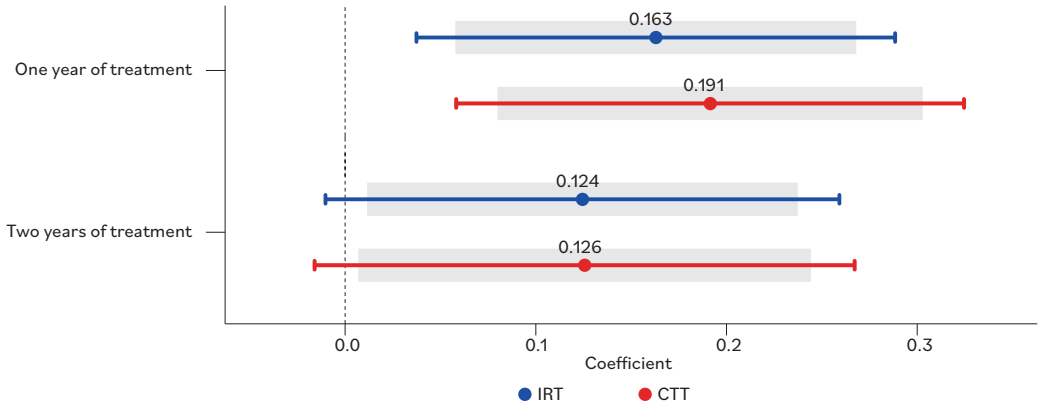
Figure 33. A visual example of the empirical strategy for science grades 3 and 4, academic year 2022-23



Source: Authors' estimations.

Overall, the curriculum reform had a positive and sizable impact on learning, equivalent to an additional 6 months of learning over the two years of the reform. The overall impact of the curriculum reform on learning is 0.142 standard deviations (SD) when using IRT scores (Figure 32). This is equivalent to a learning gain of 6 months. The effect is statistically significant for both mathematics and science assessments. The lower bound of the 95 percent confidence interval of the overall effect is 0.029 SD. A conservative reading of the magnitudes of the effect would state that the reform led to learning gains of *at least* one-tenth of a school year with 95 percent certainty. The effects of the reform

Figure 34. Impact of the curriculum reform, by years under the new curriculum



Notes: Panel sample of students. Separate regressions for each round of data collection. Because the data for the impact evaluation comes from a sample of students, we compute confidence intervals. That is, if one were to assess the total student population instead of a sample, the true parameter would lie in the range delimited by the red and blue lines with 95 percent certainty, and in the range delimited by the gray bar with 90 percent certainty.

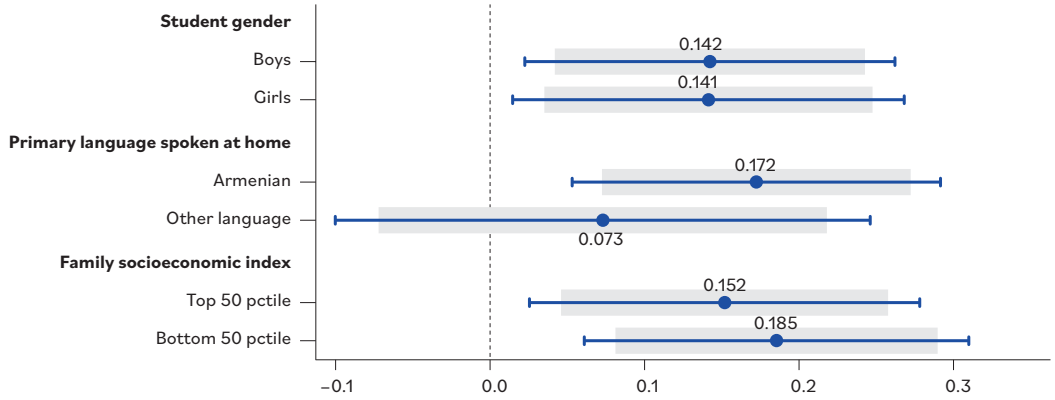
using the IRT scores are usually of lower magnitude than when using CTT scores. The impact of the reform was very similar across mathematics (0.142 SD) and science (0.141 SD). An illustration of the impact of the reform in a more didactic way clarifies these results (Figure 33). For example, when looking at the distribution of science scores in 2023 for grades 3 and 4, with the average of each group marked by a vertical line, it is apparent that the distribution of scores is similar across regions for grade 4, which used the previous curriculum. The key result lies in the average gap between grades. It is usual that students in grade 3 get lower test scores in the same assessment than their counterparts in grade 4. However, the key result is that *the average gap in knowledge between grades* is lower in Tavush – where grade 3 students used the new curriculum – than in Lori and Shirak – where they used the previous curriculum.

The impact of the reform was larger in the first year of implementation, and there was a 30 percent reduction in the magnitude of the effect by the second year (Figure 34). This reduction was due to the mathematics subject in the younger cohort and the sciences subject in the older cohort. Despite the decrease, the new curriculum still had a statistically significant impact on student performance after two years, as shown IRT scores at the 90 percent confidence level.

Was the impact of the curriculum reform different across students, teachers, and principals?

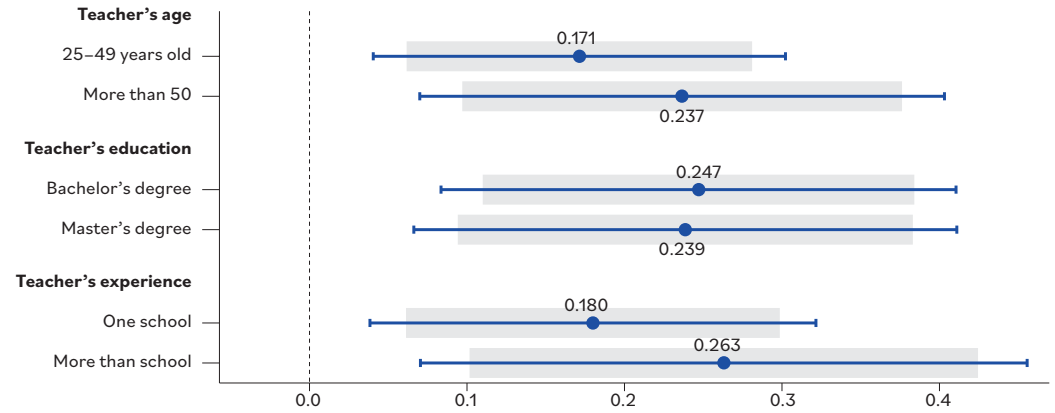
Overall, the reform did not have a differential effect across different characteristics for students, teachers, and principals across most dimensions, with a few exceptions. For example, there were no differential impacts across student gender and family socio-economic index, that is, even though the average estimated effects are higher for some groups, they are not statistically different from each other. However, the effect is clear for students whose family speaks mainly Armenian at home, but it is not statistically significant for those whose families speak another language (Figure 35). In addition, the positive impact of the reform on student learning is consistent regardless of the age, education, and experience of the teachers (Figure 36). Similar to the results across some student characteristics, the average effects are higher for some groups, but they are not statistically distinct from each other. Finally, the positive impact of the reform on student learning is consistent regardless of the education and experience of principals (Figure 37). The effect is not statistically significant for younger principals and principals with poor management practices.

Figure 35. The impact of the reform on learning was similar across student characteristics



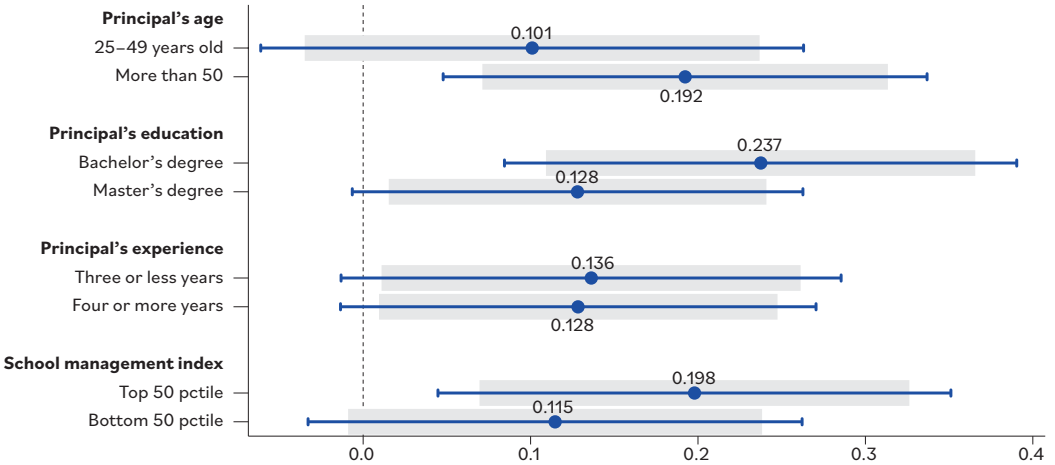
Notes: Total sample of students, appended by rounds of data collection. To do this, we construct indicator variables for different characteristics and interact this with the treatment variable.

Figure 36. The impact of the reform on learning was similar across teacher characteristics



Notes: Total sample of students that could be matched with their teachers for each subject, appended by rounds of data collection. The average impact for this subsample is 0.18SD.

Figure 37. The impact of the reform on learning was similar across principal characteristics

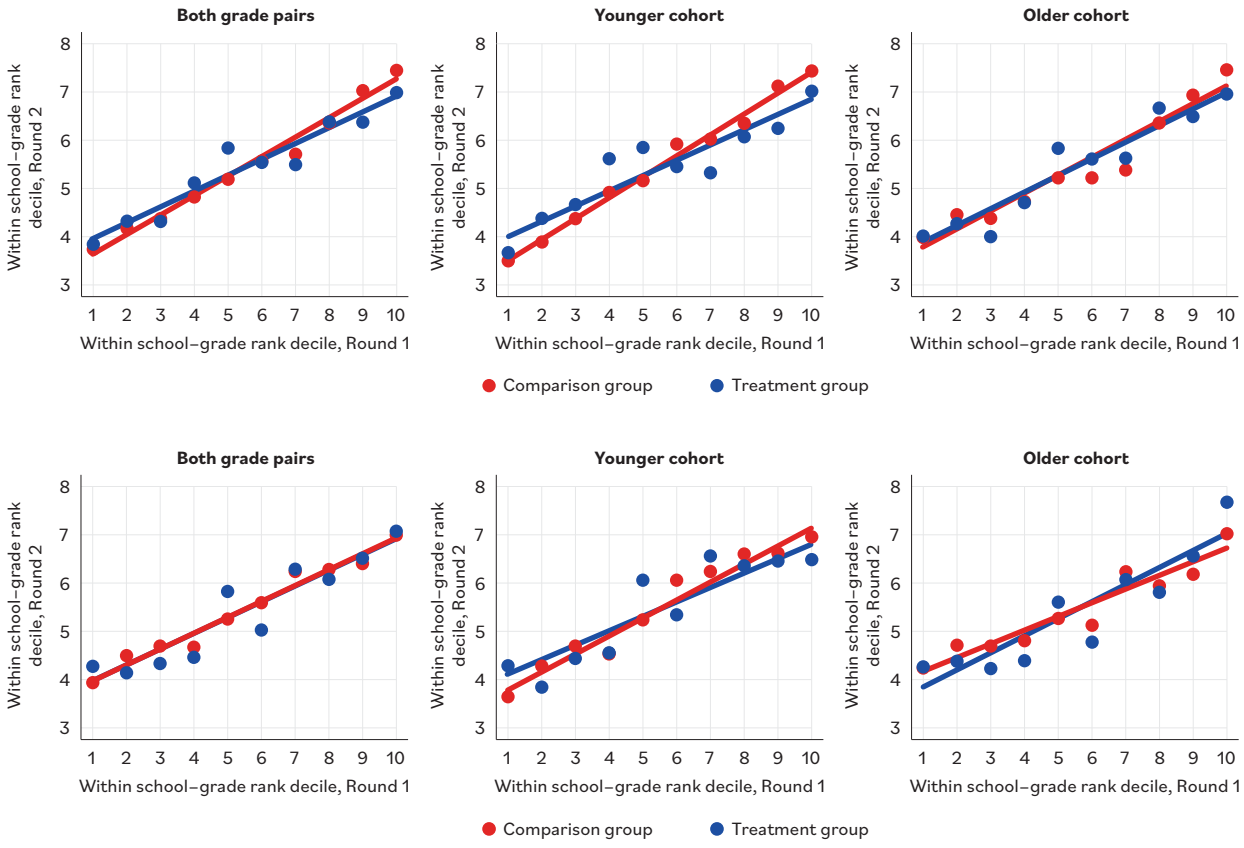


Notes: Total sample of students, appended by rounds of data collection.

Did the reform help reduce the gap in learning across low and high performing students?

The reform seems to have benefitted low-performing students at least at the same rate or more than high performing students. Prior to the reform, it was observed that teachers tended to primarily concentrate on instructing the high performers, employing a "weeding out" approach, and often holding average students responsible if they could not keep up. Within school-grade ranks of student performance in year 1 and year 2 are computed separately and correlated to investigate how student performance evolved across the entire ability spectrum from year 1 to year 2. On average, the students who were initially at the bottom of the distribution in the younger cohort of the treatment group improved more than those in the comparison group (Figure 38). This is also true in the younger cohort, but not in the older cohort.

Figure 38. Effects across the ability spectrum in Year 2, by year 1 performance



Notes: Deciles were created to compare students' performance with their peers in the two rounds of data collection for the panel sample. First, all students were ranked based on their assessment scores within their school and grade level. Then, that ranking was transformed into deciles, where the first decile represents the bottom 10 percent of the grade distribution, and the 10th decile represents the top 10 percent of performers. Students were grouped by their decile placement in the first round, and by whether they were in the treatment or comparison groups. The average decile placement in the second round was then calculated. The slope of the treatment group is flatter than the slope of the comparison group indicates that low performing students are improving and not staying behind relative to the comparison group.

8. Initial Lessons Learned and Next Steps

In 2018, Armenia embarked on an ambitious STEM curriculum reform, and while it is still underway, assessing the process to date provides initial lessons that can contribute to the global knowledge base. Armenia is implementing the reform to provide opportunities for students to reach their full potential and contribute to the social, economic, and political development of the country. These reforms reflect the Armenian government's effort to transition into a competency-based education system through more inquiry-based, student-centered, and outcome-oriented teaching, learning, and assessment. This report described four key aspects of the curriculum reform process – frameworks and design, teaching and learning materials, professional development, and learning assessment – as well as the expert working group model and stakeholder engagement that helped ensure excellence, coherence, and broad support for the reform. The report then provided assessments of how each of these aspects was implemented in the pilot rollout in the Tavush region, insights made possible by the extensive collection and use of several different types of data.

A range of observational, qualitative, and rigorously estimated evidence shows that the curriculum reform is succeeding in modernizing STEM education and increasing student learning. Despite some challenges, the pilot in Tavush has been effectively implemented, with teachers and school administrators trained and receiving ongoing mentoring, new teaching and learning materials provided, and lab environments upgraded. Importantly, a rigorous differences-in-differences analysis shows that the pilot had a positive and sizeable impact on student achievement in math and science, equivalent to an additional six months of learning over the two years of implementation. The evaluation exploits variation in the regions and grades affected by the pilot implementation following a differences-in-differences approach. The first difference is the difference in learning between consecutive pairs of grades, where one grade was not targeted by the reform and the other was. The second difference is the difference in the learning gap between pairs of grades across Tavush and comparison regions. Using the curriculum-based student assessments designed as part of the reform process, the evaluation finds that students who experienced the two-year pilot of the new curriculum made significant relative learning gains in both math and science subjects tested, equivalent to six months of additional learning. Importantly, the implementation of the curriculum reform in Tavush seems to have had a positive impact across students, teachers, and principals of different characteristics.

From Armenia's achievements, several relevant messages emerge for other countries contemplating an extensive curriculum reform, the first being the value of building local expertise and integrating stakeholder feedback for enduring capacity and support for an inevitably long and complex process. The working group model created for the curriculum revision process in Armenia involved local subject matter and education experts working on creating

the content and processes of the new curriculum with support from international experts. The local experts went from designing the curriculum, to working on the textbooks, to helping with the design and implementation of the teacher training for the pilot. Also, stakeholder feedback was regularly collected from the beginning of the process when the new curriculum subject standards were being designed, and the feedback loop was sustained as teachers were informed of what was to be expected at the beginning of the training process, and they expressed support for the reform in the focus groups organized to understand their priorities and needs.

Second, providing Teaching and Learning Materials that effectively support implementation of the new curriculum requires a realistic and well-planned timeline, including opportunity to address inevitable delivery challenges and revise drafts based on user feedback. Some materials were too complex and had to be revised based on teacher feedback before being more widely distributed. In addition, procurement processes can be very time-consuming particularly for technically specific equipment such as science labs, and schools' existing conditions may need upgrading before they can get certain equipment. These and other issues meant that teaching and learning materials were not available when the pilot started, which hindered teaching and learning. Teachers relied on other materials which included printouts of various chapters, handouts of activities for in-class instruction, and online exercises. This shows that detailed planning and tight coordination between all stakeholders clearly defining roles and timelines is critical in achieving the best results from a curriculum revision process.

Third, high-quality training to teachers and administrators to implement the new curriculum is critical, but ongoing mentoring may be equally or even more important. While there is no data yet on changes in teachers' practices, their feedback and independent assessment of recorded sessions suggests that the training for the pilot was of good quality. However, there were limited opportunities for active participation, and the mentoring program set up to help and guide teachers seems to have provided the support they needed in their day-to-day work to implement the new curriculum. The mentoring scheme also had a positive impact on teachers' new learning experience and self-confidence. Teachers found the regular peer-to-peer feedback mechanism facilitated by the mentors, where they were able to share opinions in groups discussing new instructional methods and the revised classroom assessment tools, very useful.

Fourth, a well-designed approach to student assessment helps enable curriculum reform to succeed, by supporting everything from more effective teaching in the classroom to informative systemic data for policymaking. Increased use of formative assessment in the classroom as well as better-designed summative assessments are two important parts of the new curriculum in Armenia. Transitioning from a fully summative means of assessment to an extensively formative assessment system within schools was supported by a workshop for teachers and school leaders as well as the ongoing mentoring sessions. The workshop concentrated on several methods of formative assessment as ways in which a teacher can provide continuous feedback to individual students in the classroom. Implementation went more smoothly with help from mentors and with collaboration between teachers through meetings and feedback sessions. The development of rigorous summative assessments to test student learning as part of the curriculum reform process provided a strong technical basis on which a new national student assessment framework was designed and approved at the beginning of 2024. With the implementation of this new framework, more informative, reliable, and comparable data will be available for the ongoing reform process and to guide future policymaking.

Finally, a data-driven and iterative curriculum reform process is likely to be a more successful process. In Armenia, many different types of data were used to better understand the context, identify international good practices, gather feedback on pilot implementation, and evaluate the impacts of the pilot. Crucially, these insights were fed back into the reform process and used to make improvements. For example, the overload issues identified by a quantitative comparison of the old and new curriculums as well as teacher feedback were addressed through a revision of the curriculum. Given the systemic nature of curricular reform, a process that provides a steady stream of information on all aspects of implementation, as well as flexibility to implement changes and address unanticipated challenges, is more likely to have its intended impacts.

For Armenia, the data-driven and iterative design of the curriculum reform process will serve the country well as implementation continues and reaches national scale. The information gathered and presented in this report serves as a powerful tool for MoESCS to further adjust the revised curriculum and make crucial decisions that will ensure a smoother implementation process for the entire country. For example, by continuing to monitor each subject and grade's curriculum for signs of overload, the government will be able to make revisions in a timely manner. For teachers and administrators, continuing to solicit their feedback and address their needs and concerns will help ensure that implementation of the new curriculum is not perceived or experienced as an excessive burden on their workload. By committing to such an approach of continuous improvement based on high-quality data and stakeholder engagement, the government of Armenia is ensuring that the curriculum reform will reach its full potential and equip the country's students with the knowledge and skills they need to thrive in the 21st century.

References

Atuhurra, J. & Kaffenberger, M. (2022). Measuring Education System Incoherence: Alignment of curriculum standards, examinations, and teacher instruction in Tanzania and Uganda. Journal of Educational Development, 92, 102598.

AYB Educational Foundation (2022). EU4Innovation, STEM Teacher Training in Tavush Final Report. Mimeo.

Béteille, T & Evans, D. K. (2021). Successful Teachers, Successful Students: Recruiting and supporting society's most crucial profession. World Bank: Washington, DC.

Bloom, N., Lemos, R., Sadun, R., & Van Reenen, J. (2015). Does Management Matter in Schools? The Economic Journal, 125(584), 647-674.

Cantoni, D., Chen Y., Yang D. Y., Yuchtman N., & Zhang Y. (2017). Curriculum and Ideology. Journal of Political Economy, 125(2), 338-392.

Castro Superfine, A., Marshall, A., & Kelso, C. (2015). Fidelity of Implementation: Bringing written curriculum materials into the equation. The Curriculum Journal, 26(1), 164-191.

Cheng, A. (2012). Teacher Perceptions of the Common Core State Standards. Mimeo.

Cochrane, S, & Cuevas, J. A. (2015) Teacher Perceptions of the Common Core Performance Standards. Georgia Educational Researcher, 12(1), 1-30.

Gouëdard, P., Pont, B., Hyttinen, S., & Huang, P. (2020). Curriculum Reform: A literature review to support effective implementation. OECD Education Working Papers, No. 239, OECD Publishing: Paris.

Hübner, N., Savage, C., Gräsel, C. & Wacker, A. (2021). Who Buys into Curricular Reforms and Why? Investigating predictors of reform ratings from teachers in Germany. Journal of Curriculum Studies, 53(6), 802-820.

International Labor Organisation (2020). Skills For a Greener Future: Challenges and enabling factors to achieve a just transition. ILO: Geneva.

Jackson, K., & Makarin, A. 2018. Can Online Off-the-Shelf Lessons Improve Student Outcomes? Evidence from a field experiment. American Economic Journal: Economic Policy, 10 (3), 226-54.

Kataoka, S., Vinh, L. A., Kitchlu, S., & Inoue, K. (2020). Vietnam's Human Capital: Education success and future challenges. World Bank: Washington, DC.

Kwauk, C. (2021). The Road to a Net-zero Economy Requires Building Girls' Green Skills for Green Jobs. Brookings Institute: Washington, D.C.

Leaver, C., Lemos, R., & Scur, D. (2019). Measuring and Explaining Management in Schools: New approaches using public data. Policy Research Working Paper No. 9053. World Bank: Washington, DC.

Mertler, C. (2011). Teachers' Perceptions of the Influence of No Child Left Behind on Classroom Practices. Current Issues in Education, 13(3), 1-34.

OECD (2013). Synergies for Better Learning: An international perspective on evaluation and assessment. OECD Reviews of Evaluation and Assessment in Education. OECD Publishing: Paris.

OECD (2020). Curriculum Overload: A way forward. OECD Publishing: Paris.

Palma, M. (2019). IT And Technical Training: The use of virtual training labs: Pros and cons. Mimeo.

Popova, A., Evans, D. K., Breeding M. E., & Arancibia, V. (2022) Teacher Professional Development around the World: The Gap between Evidence and Practice. The World Bank Research Observer, 37(1), 107–136.

Pritchett, L. & Beatty A. (2015). Slow Down, You're Going Too Fast: Matching curricula to student skill levels. International Journal of Educational Development, 40, 276-288.

Rodriguez-Segura, D. (2020). Strengthening Early Literacy Skills Through Social Promotion Policies? Intended and unintended consequences in Costa Rica. International Journal of Educational Development, 77, 102243.

Rodriguez-Segura, D. & Mbiti I. (2022). Back to the Basics: Curriculum reform and student learning in Tanzania. RISE Working Paper, 22/099.

Scheckler, R (2003). Virtual Labs: A substitute for traditional labs? The International Journal of Developmental Biology, 47(2-3), 231-236.

Schmidt, W., & Houang R. T. (2007). Lack of Focus in the Mathematics Curriculum: Symptom or cause. In Tom Loveless (Ed.), Lessons Learned: What International Assessments Tell Us about Math Achievement, 65-84.

Smith, J. M., & Kovacs, P. E. (2011). The Impact of Standards-based Reform on Teachers: The case of 'No Child Left Behind'. Teachers and Teaching: Theory and Practice, 17(2), 201-225.

The Government of the Republic of Armenia (2021). Government Programme 2021-2026. Yerevan: Armenia.

World Bank (2014). IT Skills Assessment in Armenia. World Bank: Washington, DC.

World Bank (2016). Needs Assessment for the Revision of Curriculum, State and Subject Standards and Syllabi within the Framework of the Education Improvement Project. World Bank: Yerevan.

World Bank (2017a). Future Armenia: Shifting to a New Growth Model: A Systematic Country Diagnostic. World Bank: Washington, DC.

World Bank (2017b). Republic of Armenia: Reducing Poverty and Improving Shared Prosperity Through Better Jobs, Skills, and Education. World Bank: Washington, DC.

World Bank (2018). World Development Report 2018: Learning to Realize Education's Promise. World Bank: Washington, DC.

World Bank (2019a). Ending Learning Poverty: What Will It Take? World Bank: Washington, DC.

World Bank (2019b). How can the World Bank help countries improve curricula to optimize learning? Thematic Group for Curriculum, Instruction, and Learning Approach Note. World Bank: Washington, DC.

World Bank (2020a). Enterprise Surveys. Armenia 2020 Country Profile. World Bank: Washington, DC.

World Bank (2020b). Realizing Armenia's High-Tech Potential. World Bank: Washington, DC.

World Bank (2021). Armenia: Teacher Profile and Policies. World Bank: Washington, DC.

World Bank (2022a). The State of Global Learning Poverty: 2022 Update. World Bank: Washington, DC.

World Bank (2022b). Armenia Human Capital Review. World Bank: Washington, DC.

World Bank (2022c). Analysis of Learning in Armenia (English). World Bank: Washington, DC.

Zacharia, Z. C., & Olympiou, G. (2011). Physical versus Virtual Manipulative Experimentation in Physics Learning. Learning and Instruction, 21, 317-331.

Annexes

Annex 1. Detailed activities of the task

As part of the ***European Commission - World Bank Partnership Programme Part III for Europe and Central Asia Programmatic Single-Donor Trust Fund - EU4Innovation STEM Pilot Activities,*** and in line with the new European Consensus on Development, the European Commission expressed an interest in ensuring that the Bank provides support to Armenia in the framework of the Trust Fund with the **specific objective** to improve the quality of STEM education in the pilot region(s) of Armenia (Tavush), as per the description below.

1. Bank-executed activities, for which the Bank has implementation responsibility:
Component A: Improving STEM education and piloting redeveloped curricula and pedagogies

Support the preparation of a work plan for the STEM education in close cooperation with European Union and MoESCS, indicating roles and responsibilities of participating stakeholders. The National Institute of Education (NIE) is expected to be a prominent stakeholder in both Bank and Recipient Executed activities.

Component A is focused on modernizing the STEM curricula and training teachers on modern, student-centred pedagogies to improve STEM education. Component A will have two subcomponents. Through this first component, the Bank will among other things, provide a consolidated draft including suggestions for the redevelopment of existing curricula (syllabi) for STEM general education (grades 1-12), including for the development of teaching and learning materials package (subject syllabi, lesson plans, draft textbooks, inputs for assessment instruments) for lower and upper secondary education (grades 5-12). The second subcomponent will focus on teacher training on modern, student-centered pedagogies and on the use of equipment and materials to be provided through the *Subcomponent C.2 Acquisition of hardware and equipment for the roll-out of the STEM Pilot*. The Bank will carry out a needs assessment of the appropriate equipment needed for the Activities based on cost effectiveness, with a view of allowing the roll out of the approach to the whole country by the government of Armenia with its own resources. The modernized curricula and materials (for grades 5-12) will support this goal and the pilot will also serve to evaluate the effectiveness of the redeveloped curricula and materials. The pilot region is Tavush and, if the funds are sufficient, it is intended to cover with the pilot partially Shirak region.

Subcomponent A.1 Redevelopment of STEM curricula and pedagogies

The redevelopment of the curricula and of the teaching and learning materials will be carried out by a working group structure led by a STEM Technical Committee (STEM TC) with support of several sub-committees. The World Bank's support under this sub-component will be provided through the advisory services and analytics for the review of STEM curriculum and preparation of terms of reference for such committees. The structure of these committees will comprise national and international education and subject area experts from the government, academia, and other relevant stakeholders who will review and produce draft revised curricula and related materials. The committees will take into consideration all previous relevant material developed in the country. Plenary meetings of the STEM TC will take place quarterly during the implementation period, with shorter virtual meetings.

The Bank's advisory services and analytics support is provided to carry out a sequence of activities including but not limited to:

a) Support for Setting up Committee Structure. Development with MoESCS of the proposal for STEME TC structures and work program.

b) Support for redevelopment of the curricula. The Bank's advisory services and analytics support will include: (i) review of the content, sequence, levels of difficulty, and coherence of the currently used STEM curricula in grades 1-12; (ii) development of inputs for teaching and learning packages, including lesson plans, draft textbooks, inputs for assessment instruments compatible with the redeveloped curricula; and (iii) teacher training modules for STEM pilot for grades 5 to 12. The pilot training module will include personality, skill development and teachers' practice for in service teachers and recent graduates with pedagogical and non-pedagogical domain expertise willing to build a career in education. The module will include the development of a leadership education program for senior school administrators to accompany the new training modules for teachers. A "hands-on" approach will be adopted for these activities to allow for the NIE to internalize the expertise and know how to replicate these activities autonomously in the future.

Subcomponent A.2 Pilot of Redeveloped STEM curricula and pedagogies

This subcomponent will carry out the evaluation of the effectiveness of new draft STEM curricula in raising aggregate learning among students in the pilot region(s). The pilot has two stages. The first stage will test individual "inputs" of the curriculum and teaching and learning materials at the classroom level. The second stage will conduct the full pilot of the redeveloped curriculum for grades 5 to 12 for one year in the pilot region(s). The subcomponent will deliver training of teachers to implement the new STEM curriculum. It will proceed to a subsequent roll out of pilot STEM teaching in lower secondary schools in the pilot region(s). The pilots will include observation of classrooms, and baseline and end line assessment of learning levels. The pilots will be done both in classroom and laboratories. It will conclude with analysis of changes to learning resulting from new practices.

The Bank's advisory services and analytics support is provided to carry out a sequence of activities including but not limited to:

a) **Teacher and senior school administration Training**—Plan and organize the development of training modules and training of STEM teachers (and school management staff) in all schools in the pilot region(s) to emphasize modern, student-centered pedagogies and innovative methodologies (blended learning) for effective delivery of STEM education based on personality, skill development and teachers' practice.

b) Development and roll-out of the pilot in the target region(s) to evaluate the new draft curricula (subject syllabi) and teaching and learning materials (lesson plans, draft textbooks, and inputs to assessment instruments).

c) Analyze the results of the STEM pilot and formulate recommendations that will be made available to the government along with the outputs (final draft curricula, draft textbooks, and associated package of materials for grades 5-12). Satisfaction surveys will be carried out with teachers and senior school administration to inform this process. A quality assurance framework will be proposed to measure impact of the whole action and will incorporate lessons learnt from the pilot programme to inform strategies of the NIE/MoESCS.

Component B: Program management, administration, and implementation support

a) Program management and administration activities for the Trust Fund, including but not limited to supporting any program governance arrangements and Trust Fund related meetings; planning and executing work plans and budgets; managing communications and conducting outreach; disseminating lessons learned; reporting on progress; and monitoring and evaluating the program.

b) Project implementation support of the Recipient-executed activities under the Trust Fund.

2. Recipient-executed activities, for which one or more Recipients (as defined in Annex 2) have implementation responsibility:

Component C: Increasing capacity of the National Institute of Education (NIE) to regulate policies affecting teaching quality and support to the STEM pilot roll-out.

This component will have two subcomponents. The first subcomponent will provide advisory services and technical assistance to raise capacity of NIE to develop and manage continuous professional development for teachers. The second subcomponent will acquire the hardware and equipment necessary for the roll-out of the STEM Pilot.

Subcomponent C.1 Increasing capacity of NIE for Professional Development of Teachers

The proposed component will work closely with Component A and will support the government in the development of the teacher professional development framework, which would outline the knowledge, practice, and professional engagement required from teachers. The development of the framework will be guided by international best practice, stakeholder consultations and alignment with the requirements set forth by the new curriculum. This component will also support the government's strategy to outsource the provision of teacher training to eligible training providers. Technical assistance will be provided to develop standards for teacher training, defining requirements for the accreditation of teacher training providers, and putting in place a proper quality assurance system to monitor quality and delivery of in-service teacher trainings.

The specific activities supported under this component, include, but may not be limited to:

a) Strengthen the Capacity for Teacher Professional Development within NIE. This activity entails targeted advisory support to guide development of teacher professional development framework, on the job training of core staff of the unit and opportunities for study tour to learn about best practices and innovations in teacher policies.

b) Supporting Outsourcing of teacher training. This activity will support development of standards for the accreditation of in-service teacher training providers and will help to set up a proper quality assurance system to evaluate and monitor the quality and delivery of teacher trainings.

The sub-component could also provide support for the development of teacher professional and subject standards. The objective of the standards would be to set clear expectations both for incumbent and existing teachers in the system in terms of what constitutes quality teaching. This activity can entail review of teacher standards in selected advanced countries to draw lessons relevant to Armenia. The work will include close coordination of work with the new curriculum revision process, which should form the basis for standard development as well as extensive stakeholder consultations.

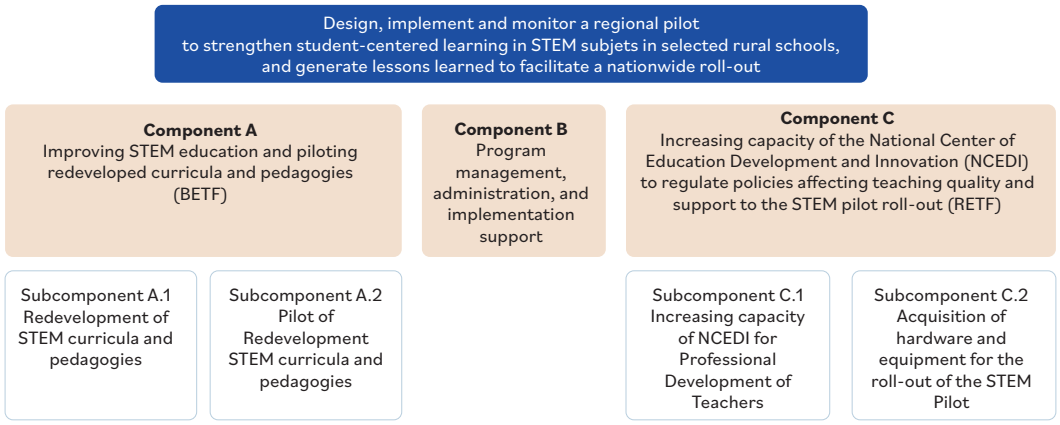
Subcomponent C.2 Acquisition of hardware and equipment for the roll-out of the STEM Pilot.

This subcomponent is strongly linked to the Component A and will acquire the hardware and material required for the successful roll-out of the STEM pilot. The subcomponent will prioritize purchase of efficient and low-cost equipment for the pilot roll-out to allow replication of the same model country wide via government financing. The sub-component will also support minor works where necessary to install new equipment and ensure the needed safety.

The overview and expected results of the task can be seen below.

Box 4. Overview and Expected Results of the Task

Overview



Expected Results

1. Enhancement of learning and teaching materials/methodologies to ensure quality education in STEM

   a) Draft STEM curricula for grades 1-12 and full package of teaching and learning materials (lesson plans, draft textbooks, and inputs for assessment)
   b) Increased capacity of NIE to ensure quality control of professional development of teachers
   c) STEM teachers and senior school administrators in pilot region(s) trained and tailor-made support provided (including in redeveloped curriculum of STEM subjects, modern student-centered pedagogies, blended learning etc.)

2. Improved learning outcomes among students

   a) Pilot region's high schools and middle schools equipped with low-cost STEM equipment to roll-out enhanced learning

Annex 2. A pathway for curriculum reform

1. **Analyze current teaching practices and learning goals**
   - i) Things that need to be changed about student learning in the existing curricula
   - ii) Recurring challenges or issues that need addressing through an adjustment in your teaching methods or curriculum.
   - iii) Things to preserve in the current curriculum
   - iv) What does student learning or student understanding mean in the context of your curriculum?

2. **Brainstorm the Ideal Graduates**
   - i) Imagine the ideal graduates and backward map from that
   - ii) Subjects, content, attitudes, habits, skills can then be defined with reference to a clear end goal

3. **Examine the links between goals and course design**
   - i) The system's goals for student learning – core disciplinary ideas, cross-cutting concepts, learning goals
   - ii) How these goals should inform the teaching approaches (active learning, student centered teaching, formative/summative assessment balance etc.)
   - iii) How does your curriculum design most emphasize the learning goals?

4. **Emphasize the role of assessment.**
   - i) Ensure students meet stated learning targets.
   - ii) Strategize the types of assessments to use.
   - iii) Map the assignments and exercises directly to the goals you have set out for students - What competencies or skills does each assignment address?

5. **Develop relevant teaching strategies and approaches.**
   - i) Active learning methods, traditional teaching approaches and their balance
   - ii) How technology might support the learning process
   - iii) Use of group work or peer collaboration assignments

6. **Gather Data**
   - i) How are the educational outcomes for students in the current system? National and international assessments

7. **Formulate, Deliberate, and Assess**
   - i) Debate is important, but the period for debate should be limited.
   - ii) Good curricula match available time with best learning modality on for specific subjects in ordered sequences that best inculcate a defined set of Habits, Attitudes, Knowledge, and Skills.
   - iii) Bad curricula are a series of unconnected details; Good curricula are a sequence and progression of important key underlying concepts.

Funded by
the European Union

THE WORLD BANK
IBRD • IDA | WORLD BANK GROUP