McKinsey&Company

Drivers of Student Performance: Insights from North America

Education 2017

Authored by:
Jake Bryant | Emma Dorn | Paul Kihn
Marc Krawitz | Mona Mourshed and
Jimmy Sarakatsannis

acknowledgements

The authors deeply thank the many people who supported us in bringing this report to fruition. We are grateful for the invaluable guidance of our analytics leadership: Rafiq Ajani, Taras Gorishnyy, and Sacha Litman. We thank our dedicated engineer and data scientist colleagues: April Cheng, Sujatha Duraikkannan, Roma Koulikov, Devyani Sharma, and Avan Vora; and we are grateful for the substantial contributions from our colleagues Anne-Marie Frassica, Joy Lim, Esteban Loria, Miriam Owens, Corinne Spears, Amy Tang, Rachel Valentino, and Paul Yuan. We further acknowledge the external thought-leaders and experts who provided counsel and expertise. Finally, this report would not have been published without the support of our editor Cait Murphy, the design creativity of Nicholas Dehaney at Spicegrove Creative.

foreword

The PISA database is open to the world – a powerful resource for everyone who wants to understand education and help improve it, from policy-makers and researchers to school leaders, teachers and parents. Over recent years, McKinsey have done just this, drawing on PISA to identify the policies and practices that make a real difference. Their work began with ground-breaking reports on The World's Best School Systems And How To Build Them. And these new regional analyses of student-level performance represent another significant milestone.

The reports suggest that students' attitudes and motivation are critical drivers of achievement. So too are their experience in the classroom, of both teaching strategies and digital technology, as well as the time they spend in education. McKinsey's perceptive insights will encourage schools around the world to discover new ways to nurture and inspire their students.

What sets these reports apart is their regional focus. I often hear countries say that learning from the world's outstanding systems is vital, but that just as powerful is the chance to learn from their own neighbours, with similar cultural backgrounds and with shared problems and opportunities.

In every country, the search is on for ways to take education to the next level, to prepare young people for a dramatic and challenging century. This is complex work. What is the right mix of policies, implementation strategies and enabling conditions – in each country and region? How should they be prioritised, sequenced and linked? If we are really to secure achievement, well being and equity, on a global basis, then these will be the issues that educators need to work on. The new reports from McKinsey offer us a fresh and welcome perspective.

Andreas Schleicher
Director for the Directorate of Education and Skills | OECD

executive summary

Policy makers, educators, and parents across North America want to raise a generation of students who can thrive amid the relentless change wrought by technology and globalization. Yet improving educational outcomes has proved elusive. Some countries, states, and municipalities have made great strides, but many continue to struggle. Educators continue to debate what matters and what works.

In this report, we take a data-driven approach to consider a few of the most active debates in North America today: around mindsets, teaching practices, the length of the school day, and early childhood education. Our data comes from the Programme for International Student Assessment (PISA), administered by the Organisation for Economic Co-operation and Development (OECD). Since 2000, the OECD has regularly tested 15-year-olds around the world on mathematics, reading, and science. The most recent assessment in 2015 covered more than half a million students across 72 countries including nearly 30,000 students in North America. What makes PISA particularly powerful is that it goes beyond the numbers, asking students, principals, teachers, and parents a series of questions about their attitudes, behaviors, and resources.

The report's findings include the following four insights:

Student mindsets have triple the effect of socioeconomic background on student outcomes.

It is hardly news that students' attitudes and beliefs—what we term their "mindsets"—influence their academic performance. But how much? And which mindsets matter most?

By analyzing the PISA data, we found that in North America mindset factors explain a greater proportion of a student's PISA score (at 37 percent) than even the home environment and student socioeconomic factors combined (at 12 percent), controlling for interactions between the variables.

Some mindsets are more important than others. In the 2015 PISA assessment, the most predictive mindset is the ability to identify what motivation looks like in day-to-day life (including doing more than expected and working on tasks until everything is perfect). We call this "motivation calibration," as it involves a student "calibrating" what types of behaviors motivated students exhibit.

Other mindsets that are predictive of student outcomes include having low test anxiety, and believing that one's school science work will be useful for one's future career.

We also found that students with a strong growth mindset (those who believe they can succeed if they work hard) outperform students with a fixed mindset (those who believe that their capabilities are static) by 16 percent.

To be clear, mindsets alone cannot overcome economic and social barriers, and researchers debate the extent to which parental or school-system-level interventions can shift student mindsets. Our research does, however, suggest that mindsets matter a great deal, particularly for those living in the most challenging circumstances.

Students who receive the right blend of inquiry-based and teacher-directed instruction have the best outcomes.

High-performing and fast-improving school systems require high-quality instruction. We evaluated two types of science instruction to understand how different teaching styles affect student outcomes. The first is "teacher-directed instruction", in which the teacher explains and demonstrates scientific ideas, discusses questions, and leads classroom discussions. The second is "inquiry-based teaching", which includes a diverse range of practices from conducting practical experiments, to understanding how science can be applied in real life, to encouraging students to create their own questions.

Our research found that student outcomes are highest with a combination of teacher-directed instruction in most to all classes and inquiry-based teaching in some classes. If all students experienced this rebalanced blend of instruction weighted in favor of teacher-led, average PISA scores in North America would be 4.4 percent (or 22 PISA points) higher, equivalent to more than half a school year of learning. Currently almost half of North American students are receiving too little teacher-directed instruction.

It's also important to note, moreover, that some kinds of inquiry-based teaching are better than others. More structured inquiry-based activities yield higher PISA scores. Understanding how a science concept can be applied, and conducting and drawing conclusions from scientific experiments, improve scores significantly. Less structured methods of inquiry, however, such as allowing students to design their own experiments, result in lower scores across the board.

Given the strong conventional support for inquiry-based pedagogy, this seems counterintuitive. We offer two hypotheses. First, students cannot progress to inquiry-based methods without a strong foundation of knowledge, gained through teacher-directed instruction. Second, inquiry-based teaching is inherently more challenging to deliver, and teachers who attempt it without sufficient training and support will struggle. Better teacher training, high-quality lesson plans, and school-based instructional leadership can

In poorly performing schools having a well-calibrated motivation mindset is equivalent to vaulting from the lowest to the highest socioeconomic quartile

help; so can giving principals and teachers the confidence to focus on fewer incidences of well-planned inquiry, rather than trying to use these methods exclusively.

While teacher-directed instruction has the most positive impact on PISA scores, inquiry-based practices do better in promoting students' 'joy in science' and instilling the belief that doing well in school will help them have a brighter future. We believe that is why blending teacher-directed instruction with inquiry-based teaching produces the greatest overall benefit across North America.

Increasing the school day can improve student outcomes, but significant gains can also be made from using existing time better.

Across the United States (the sole country focus for this finding) federal, state, and district officials have been considering the merit of longer school days, pushed to some extent by advocacy groups such as the National Center on Time & Learning.

According to the PISA data, student outcomes improve by 3.9 percent for each additional half-hour of instruction, up to 6.5 hours per day. If all students performed at the level of those currently receiving 6 to 6.5 hours of instruction per day, this would boost average science achievement by approximately 4.4 percent (or 22 PISA points). But extending the school day is expensive, and has knock-on effects on commutes, after-school activities, and school start-times. It's also worth noting that some European systems—such as those in Finland and Germany—achieve superior results to the United States despite having a shorter average school day. This suggests that lengthening the school day is only part of the answer. Improving the quality of every hour in school also remains critical.

Early childhood education has a positive academic impact in most regions, but not in the United States.

Many studies have shown that high-quality early childhood education (ECE) improves social and academic outcomes, producing gains that sustain for many years. But research also shows that low-quality early childhood programs are less effective, with benefits fading out over time.

Our findings suggest that ECE is indeed having a positive impact on outcomes at age 15 in almost every region in the world. For example, students who had some ECE score 19–21 PISA points higher in Asia and Europe, controlling for student socioeconomic status, school type, and location. Canada reflects this global trend with a 15 PISA point (or 3 percent) lift for children who received ECE.

In the United States, however, students with ECE actually score 8 PISA points (or 2 percent) lower than those without. What is going on with ECE in the United States? While the PISA data cannot provide an answer, it does provide hints.

For example, it seems that lower socioeconomic status children are the ones who are benefiting the least. It also seems that children who start ECE at age one or younger score dramatically lower on PISA than those who start later. This suggests that there could be a significant ECE quality problem in the United States for low-income and very young children, potentially linked to the lack of government provision compared with many other developed nations.

As we share these four findings, we are mindful of their limits. One cannot construct definitive answers from a single source, no matter how broad or well designed. The direction of causality, sample sizes, missing variables, and nonlinear relationships are all potential issues. Many questions still need to be resolved through a thoughtful research agenda and longitudinal experimentation. That said, we believe that these four findings provide important insights into how students succeed—and that North American educators should heed these lessons as they develop learning agendas and school improvement programs to deliver the progress that their students deserve □

introduction

Effective education is essential for children as they journey towards constructive citizenship and productive lives, and as societies address inequality and face the heady change of uncertain times. No wonder, then, that there is broad interest in understanding how to create high-performing school systems that serve everyone well, regardless of background, and how to improve those systems that are not making the grade.

For the past decade, we have studied these issues. In 2007, McKinsey's *World's Best Performing School Systems* report explored why some school systems consistently perform better than others. The report highlighted the importance of getting the right people to become teachers, developing their skills, and ensuring that the system is able to deliver the best-possible instruction for every child. In 2010, our *World's Most Improved School Systems* report explored what it takes to achieve significant, sustained, and widespread improvements in school system performance: from poor-to-fair-to-good-to-great-to-excellent.[1]

These two reports focused on interventions at the system level. Now, in this current report, we undertake a quantitative analysis at the student level. To do so, we apply advanced analytics and machine learning to develop insights from the world's deepest and broadest education data set, the Programme for International Student Assessment (PISA), run by the Organisation for Economic Co-operation and Development (OECD).

Begun in 2000 and repeated every three years since, PISA examines 15-year-olds on applied mathematics, reading, and science. The most recent assessment covered nearly 540,000 students in 72 countries.[2] These numbers include 30,000 students in the United States and Canada, which constitute the North American region under discussion here. (We have included Mexico in the Latin American region). PISA

EXHIBIT 01: PISA IS A RICH SET OF ASSESSMENT AND SURVEY DATA.

OECD PISA test performance + survey data

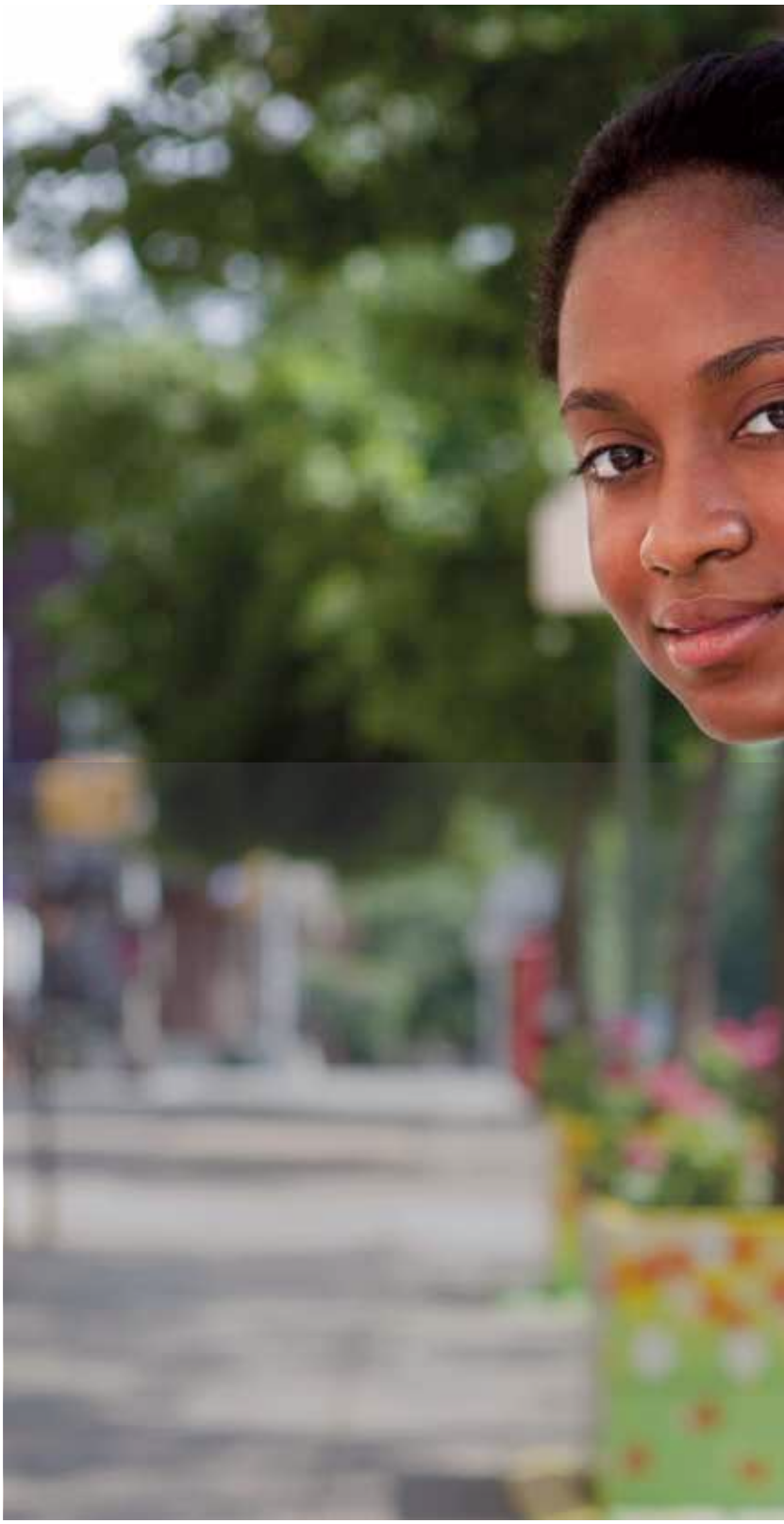
72	18,000	140,000	110,000	540,000
countries[1]	schools	parents	teachers	students
3	~270	~150	~250	~770
subjects	school variables	parent variables	teacher variables	student variables
• Math • Science • Reading	E.g., • Size • Resources • Governance and autonomy • Extra-curriculars	E.g., • Education • Income • Employment • Attitudes to school and education	E.g., • Experience • Certification • Professional development • Teaching strategies • Assessment strategies	E.g., • Attitude to study and learning • Growth mindset • Problem solving approach • Repeated grade • Economic and social status

1 Report analysis excludes Albania as it was not possible to match test and survey data; includes Argentina, Kazakhstan, and Malaysia despite sampling concerns as our analysis examines drivers at the student level rather than country-level comparisons.

test-takers also answer a rich set of attitudinal questions. Students, teachers, parents, and principals completed surveys that provide information on home environment, economic status, student mindsets and behaviors, resources, school leadership, teaching practices, teacher background, and professional development (Exhibit 1). The 2015 PISA focused on scientific performance. The survey questions therefore largely addressed science teaching and learning. Half of the student assessment was related to science, with the other half split between reading and math.[3]

Standardized tests have their shortcomings. They cannot measure important soft skills or non-academic outcomes, and they can be subject to teaching-to-the-test and other "gaming" of the system. Even so, we believe that PISA provides powerful insights on global student performance, especially because it aims to test the understanding and application of ideas, rather than facts derived from rote memorization.

In this report, we examine educational performance in North America (the United States and Canada). Canada consistently scores in the top ten countries globally in science, reading, and mathematics, but since 2006, its PISA scores have been flat or fallen slightly. In the United States, 2015 PISA scores were just above the OECD average in science and reading, but below in mathematics. Scores have not budged much since 2006, but equity has improved due to gains in performance among disadvantaged students.[4]



In our analysis, we look beyond the macro-performance of each country to examine the factors driving individual student performance. While we focus on the 2015 PISA results, we also consider previous test administrations, using a range of traditional and advanced analytical techniques. First, we used SparkBeyond, a supervised machine-learning and feature-discovery tool that identified variables and groups of variables that were most predictive of student performance. We then applied more traditional descriptive and statistical analyses to those features that were shown to be most important in contributing to students' PISA performance. (For more, see the Analytical Appendix at the end of the report.)

We also examine how patterns differ by the school performance levels outlined in our 2010 report (poor, fair, good, great, excellent), and by student economic, social, and cultural status (or ESCS—see the Analytical Appendix for a detailed explanation). Our research resulted in four key findings for North America, regarding mindsets, teaching practices, duration of classroom instruction, and early childhood education. These findings emerged as both highly predictive of student performance and potentially responsive to school system interventions. They should therefore be the subject of further exploration by policymakers and practitioners □

Finding 1: Student mindsets have triple the effect of socioeconomic background on student outcomes

The role of mindsets in educational achievement is a nascent but intriguing field of study. In her 2006 book, *Mindset: The New Psychology of Success*, Carol Dweck argued that individuals with "growth mindsets"—that is, those who believed that their success was due to hard work and learning—were more resilient and likely to be motivated to succeed than those with "fixed mindsets"—those who believed that their innate abilities were static and could not be developed. Dweck also argued that growth mindsets could be taught. A large-scale 2016 Stanford study of all tenth graders in Chile—the largest to date—found that having a strong growth mindset rivals socioeconomic status in predicting achievement, and that low-income students with strong growth mindsets were able to achieve at the same level as high-income students with fixed mindsets.[5]

In 2016, Angela Duckworth highlighted the importance of another mindset trait—"grit"—as a predictor of performance in *Grit: The Power of Passion and Perseverance*.[6] Researchers have also explored the role of broader character traits such as perseverance, curiosity, conscientiousness, optimism, and self-control in children's success. Others, however, have questioned both the magnitude of the effect, and the usefulness of interventions in this area.[7]

We had three objectives in analyzing the mindset data available in the PISA dataset: to quantify the impact of mindsets on student performance; to assess which mindsets matter most; and to understand which types of schools and students benefit the most from certain mindsets. To quantify the impact of mindsets, we sorted the 100 most predictive variables (see the Analytical Appendix for more detail) emerging from the PISA surveys into a number of specific categories: mindset factors, home environment (including socioeconomic status), school factors, teacher factors, student behaviors, and others. We then separated mindsets into two types: "subject orientation" and "general mindsets." Subject orientation refers to a student's attitudes about science as a discipline (science, specifically, because that was the focus of the 2015 PISA). General mindsets refer to a student's broader sense of belonging, motivation, and expectations.

To be conservative, we excluded from the analysis variables where secondary research suggests the direction of causality is largely from score-to-mindset rather than from mindset-to-score. For example, we judged that students' academic performance is more likely to influence their future educational expectations (whether they will complete college) than the other way around, and thus excluded this variable from our model.

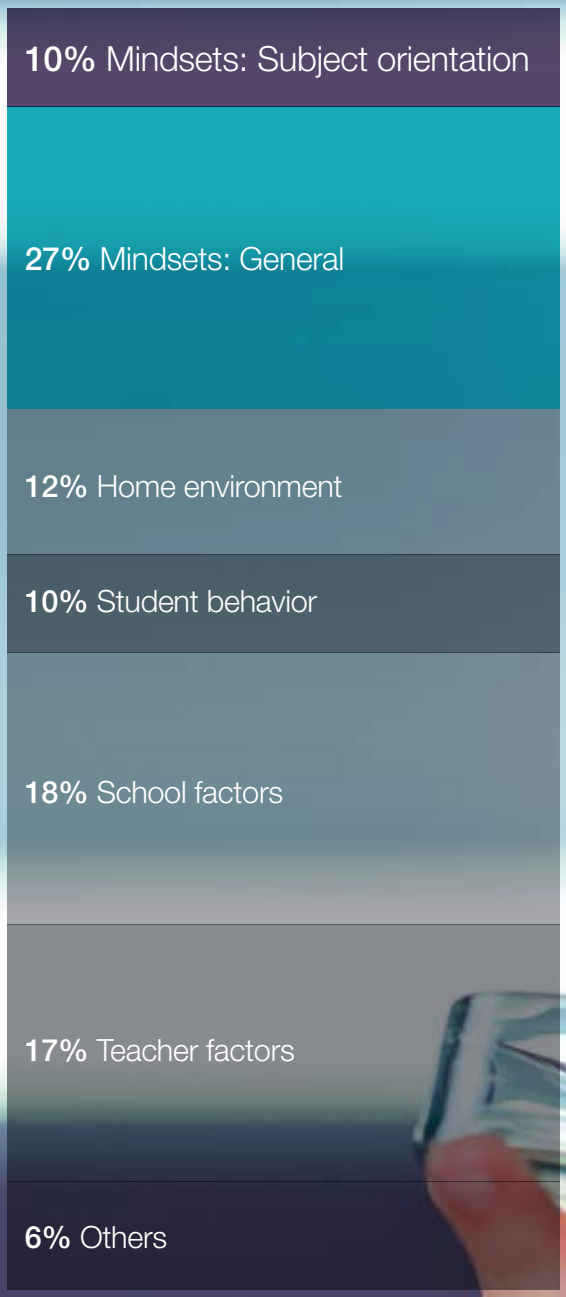
We then determined how influential each category was in terms of predicting student performance. Our conclusion: controlling for all other factors, student mindsets are three times as powerful (at 37 percent of total predictive power) as the home environment (12 percent), and as much as all school and teacher factors combined (Exhibit 2).[8] Furthermore, general mindsets account for almost three-quarters of the effect found. In all five regions of the world that we studied[9], student mindsets have more influence on student PISA score than home environment, reinforcing the prevalence and potential importance of this finding.

This finding may seem surprising in light of previous research that has highlighted the importance of the mother's education level[10] and of the quality of teachers.[11] This is largely because of what each researcher measured. Previous research on the role of maternal education did not query or control for student mindsets, so this effect was not uncovered. Meanwhile, survey tools, such as PISA, are inherently unable to fully measure the quality of teachers—thus teacher factors are likely under-represented in our analysis. What we can say is this: demographics and teacher quality are still critical, but we now know that mindsets matter too.

EXHIBIT 2: MINDSETS ECLIPSE HOME ENVIRONMENT IN PREDICTING STUDENT ACHIEVEMENT.

FACTORS DRIVING NORTH AMERICAN STUDENT OECD PISA SCIENCE PERFORMANCE 2015

% of predictive power by category of variable



Examples of subject orientation mindsets:

"I have fun learning science"

"I am interested in the universe and its history"

"Air pollution will get worse over the next 20 years"

Examples of general mindsets:

"I see myself as an ambitious person"

"What I learn in school will help get me a job"

"I feel like I belong at school"

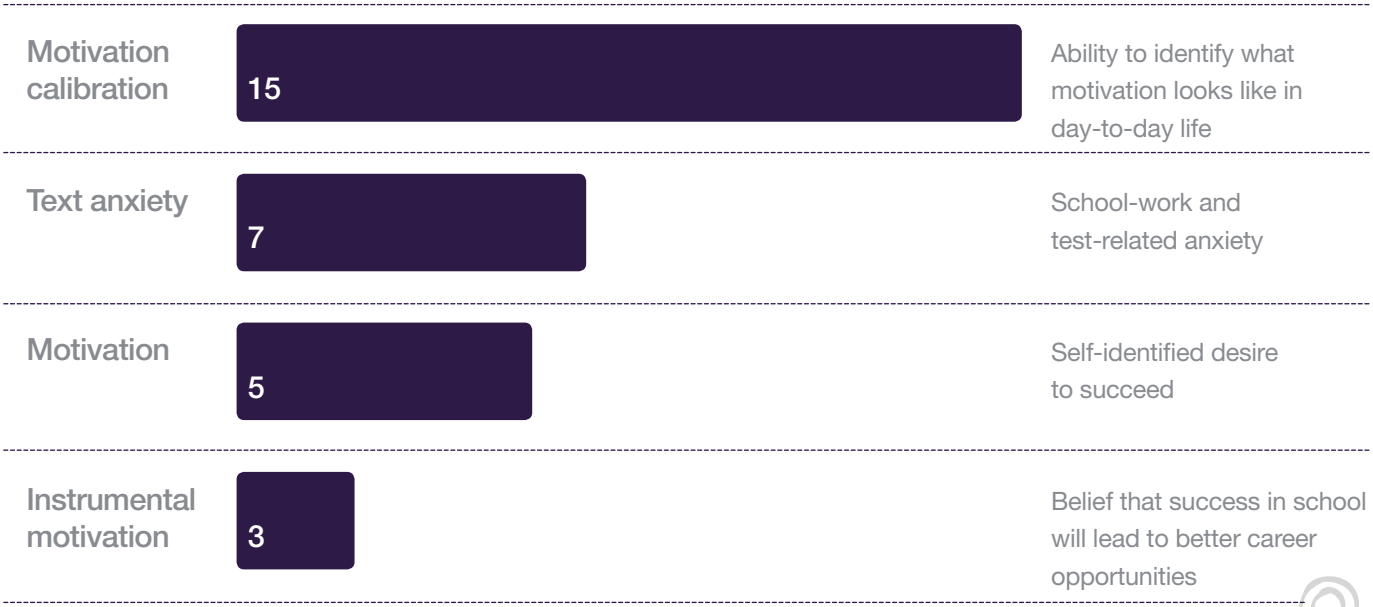
"If I put in enough effort, I can succeed"

Source: OECD PISA 2015, McKinsey analysis

EXHIBIT 3: WHAT MINDSETS MATTER MOST?

Score improvement for top 2015 general mindset measures[1]

Percent increase in PISA Science score



1 Statistically significant in regressions controlling for student socioeconomic status, school type, and school location.
Source: OECD PISA 2012, 2015; McKinsey analysis

Our analysis also found that some specific mindsets are more important than others in improving student outcomes (Exhibit 3).

"Motivation calibration"—defined as the ability of students to correctly assess what motivation looks like—was the single most predictive mindset. To measure this, in 2015 PISA asked test-takers to assess the motivation of three hypothetical students (Exhibit 4).

EXHIBIT 4: WHAT IS MOTIVATION CALIBRATION?

Student evaluation of the motivation of other students: "Is the following student motivated?"

Sienna gives up easily when confronted with a problem and is often unprepared for class.

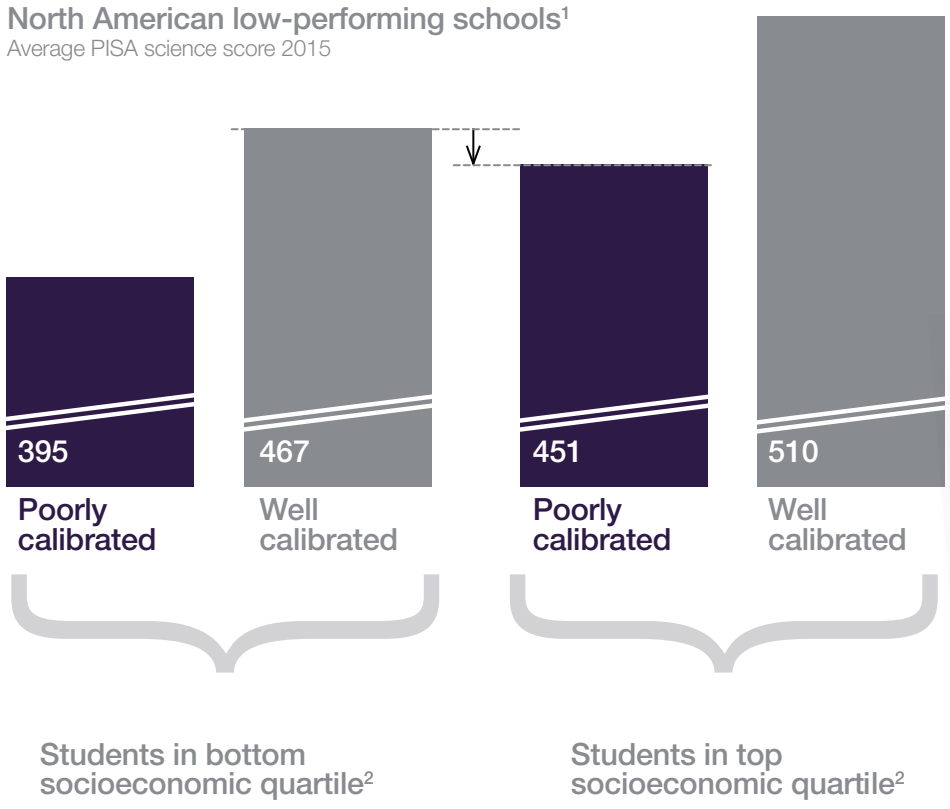
Jed mostly remains interested in the tasks he starts and sometimes does more than what is expected of him.

Jemina wants to get top grades at school and continues working on tasks until everythingis perfect.

Source: OECD PISA 2015, McKinsey analysis

Based on the responses to these questions, we created an index of motivation calibration (see the Analytical Appendix). What we found was that simply understanding what motivation looks like in daily practice is a powerful performance indicator. Students who have good motivation calibration score 15 percent (or 67 PISA points) higher than poorly calibrated students. This relationship holds even after controlling for socioeconomic status, location, and type of school. In contrast, students who self-identify as "wanting to be the best and wanting top grades" score just five percent higher than those who do not. Why is this the case? Our hypothesis is that knowing what it actually takes to succeed is more important than just wanting to be the best. Furthermore, students are more likely to be honest when talking about a third person versus directly assessing their own

EXHIBIT 5: HAVING A WELL-CALIBRATED MOTIVATION MINDSET ALLOWS LOW SOCIOECONOMIC STUDENTS TO OUT-PERFORM THEIR POORLY CALIBRATED, HIGH SOCIOECONOMIC STATUS PEERS



1 Schools with average PISA score of under 480. These schools serve 37% of North American students.
2 Using PISA's index for economic, social, and cultural status (ESCS) as a proxy for socioeconomic status; statistically significant in a regression controlling for public/private and urban/rural school types.
Source: OECD 2015 PISA, McKinsey analysis

motivation, and calibration itself is actually important. Students cannot exhibit positive study behaviors if they do not know what they look like.

Interestingly, the relationship between motivation calibration and PISA scores is strongest for students in low-performing schools. In fact, for those in low-performing schools, having a well-calibrated motivation mindset is equivalent to vaulting into a higher socioeconomic status. Students in the lowest socioeconomic status quartile who are well calibrated perform as well as those in the highest socioeconomic status quartile who are poorly calibrated (Exhibit 5).

These findings are consistent with those of previous PISA tests. In 2012, for example, PISA asked about growth versus fixed mindsets. (They did not repeat these questions in 2015). Specifically, students answered questions about the extent to which they agreed that their academic results where fixed ("I do badly whether or not I study") or could be changed through personal effort ("If I put in enough effort, I can succeed" or "If

I wanted to, I could do well"). Students with a strong growth mindset outperformed students with a fixed mindset by 16 percent (or 71 PISA points) in North America. The effect of growth mindsets parallels what we observed about motivation calibration in the 2015 data: in 2012, within low-performing schools, students from the lowest socioeconomic quartile but with a strong growth mindset outperform by 12 percent students from the highest socioeconomic quartile who have a fixed mindset.

Unfortunately, the students who would most benefit from these mindsets are least likely to have them. For example, 67 percent of low-socioeconomic-status students demonstrate strong motivation calibration, compared with 80 percent of high-socioeconomic-status students. Similarly, only 24 percent of low-socioeconomic-status students have a strong growth mindset, compared with 38 percent of high-socioeconomic-status students. This reflects systemic factors that go well beyond the influence of schools, but also underlines the potential benefit of mindset-oriented approaches.

To determine the power of a system-level intervention, we investigated how scores might change if different mindsets could be shifted across the population and if the relationship between mindset and score held steady. Our analysis suggests that shifting all students to a well-calibrated motivation mindset would result in a 3.6 percent overall score improvement across North America. Shifting all students to a growth mindset could improve scores by 3.7 percent.

To be clear, mindsets alone cannot overcome economic and social barriers. This research does suggest, however, that they are a powerful predictor of student outcomes, particularly for those living in the most challenging circumstances. The question is what, if anything, can be done to improve mindsets at a system-wide level. Research is being done to answer that question—much of it focused on North America students—and there are promising indications that it may be possible for schools to make effective interventions.

For example, on growth mindsets, a 2015 study of 1,500 secondary school students in 13 different schools, rich and poor, from all over the United States, found that growth-mindset and sense-of-purpose interventions delivered significant results. The researchers administered two 45-minute online modules to students over the course of a semester. The growth-mindset modules provided direct instruction on the physiological growth potential of the brain given hard work. The results were positive: students at risk of dropping out of high school, constituting a third of the sample, increased their grade-point averages (GPA) in core academic courses by 0.13 to 0.18 (on a 4.0 scale), and their core course pass rates increased by 6.4 percent.[12]

Similarly, on motivation calibration, recent research suggests that meta-cognition and self-regulation strategies can improve student outcomes. Interventions to help students plan, monitor, and evaluate their learning may be a promising way to improve student motivation and perseverance as they tackle challenging academic content.[13]

Such research is a work in progress, but these and other experiments indicate that harnessing the power of mindsets may be a promising way to support achievement—in addition, of course, to teaching fundamental academic content □

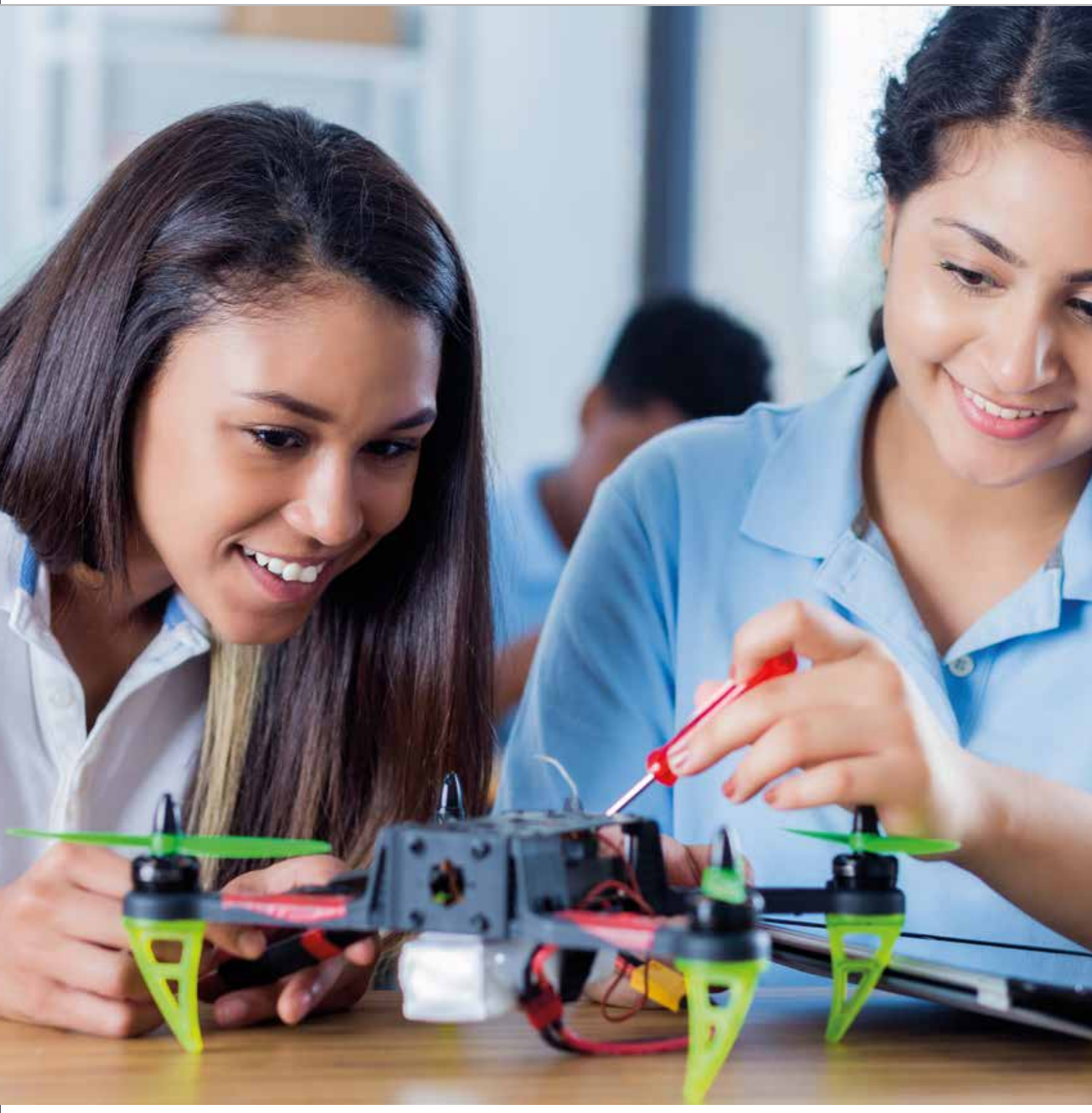
Girls and Science

Why don't more girls embrace careers in science, technology, engineering, and mathematics (STEM)? Educators, policy makers, and business leaders are all trying to encourage more women to consider STEM jobs, but the gender gap remains significant. The PISA data can help us understand why, and what might be done.

In Canada, there is no statistically significant difference between the PISA science scores of girls and boys. In the United States, boys score on average seven PISA points higher than girls. Across all OECD countries, boys score just four points (less than 1 percent) higher. There are some more subtle differences, however. For example, boys are more than twice as likely as girls to expect to work as engineers, scientists, or architects; and girls are more than three times as likely to expect to work in health professions.[14]

Looking specifically at the predictive mindsets highlighted in this report, we find that girls in North America are more likely than boys to have strong motivation calibration and self-identified motivation, and to believe that their school science work will be useful in the future. But girls are also more likely to have schoolwork- and test-related anxiety. While 45 percent of North American boys say they experience test-related anxiety, 69 percent of girls do—the biggest gap of any region. Girls are also less likely to have 'joy in science' (defined as agreeing with statements such as "I have fun when I am learning science" and "I like reading about science"). While 65 percent of North American boys have high joy in science, only 57 percent of girls do. Our analysis suggests that girls' anxiety and lack of joy in science, in effect, might be cancelling out their higher motivation calibration and affecting the choices they make.

The PISA data proves that girls have the building blocks, in terms of academic outcomes and positive motivation mindsets, for STEM careers. If interventions were made to decrease their sense of anxiety and increase their enjoyment of science, the outcome might well be more female STEM professionals.



Finding 2: Students who receive the right blend of inquiry based and teacher-directed instruction have the best outcomes

Teachers matter. Multiple research reports, including our own, have demonstrated that high-performing school systems require effective teachers and teaching. The challenge, then, is to determine what teaching practices work, and how teachers can deliver high-quality instruction.

We evaluated two types of science instruction to understand the relationship between teaching methods and student outcomes. The first is "teacher-directed instruction" where the teacher explains and demonstrates scientific ideas, discusses student questions, and leads class discussions. The second is "inquiry-based teaching," where students play a more active role. Inquiry-based teaching spans a diverse range of practices, from conducting and drawing conclusions from practical experiments, to understanding how science can be applied in real-life, though to more unstructured activities such as encouraging students to create their own questions, design experiments to test their hypotheses, and argue about science questions (Exhibit 6). There is active debate over which approach is preferable and which practices lead to better student outcomes.[15]

EXHIBIT 6: OECD PISA ASKED STUDENTS HOW OFTEN THEY EXPERIENCED THE FOLLOWING TEACHING PRACTICES.

How often does this happen in your school science class…

Teacher-directed instruction

- The teacher explains scientific ideas.
- A whole class discussion takes place with the teacher.
- The teacher discusses our questions.
- The teacher demonstrates an idea.

Inquiry-based teaching

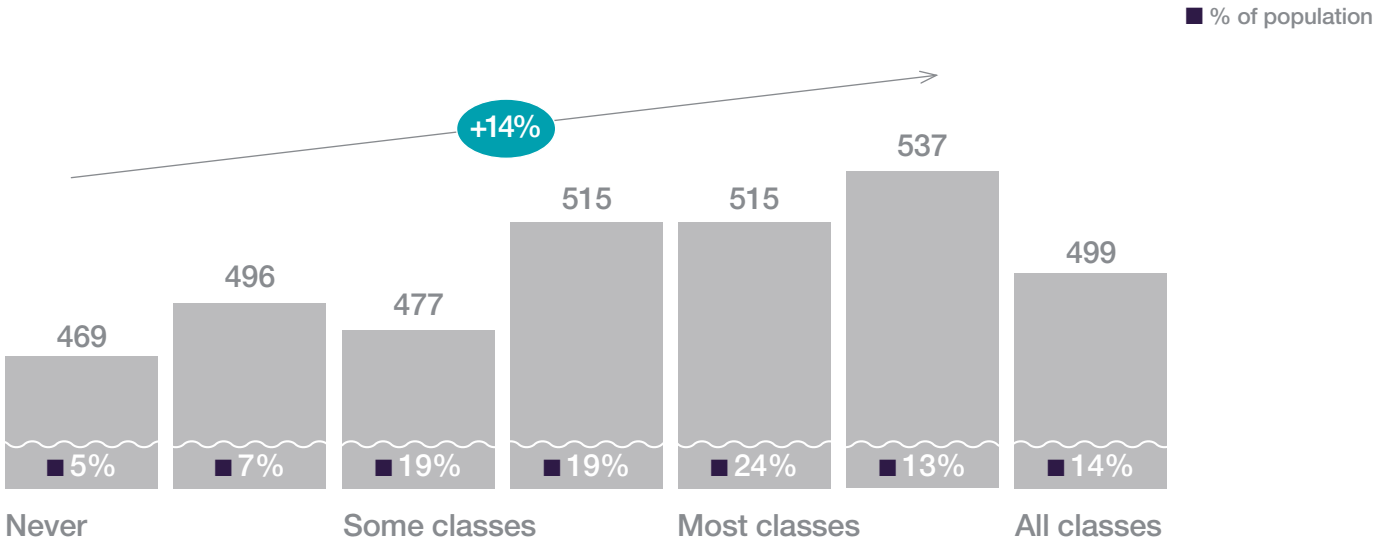
- Students are given opportunities to explain their ideas.
- Students spend time in the laboratory doing practical experiments.
- Students are required to argue about science questions.
- Students are asked to draw conclusions from an experiment
- The teacher explains how science ideas can be applied
- Students are allowed to design their own experiments.
- There is a class debate about investigations.
- The teacher explains the relevance of concepts to our lives.
- Students are asked to do an investigation to test ideas.

Source: OECD PISA 2015

EXHIBIT 7: WHEN TEACHERS TAKE THE LEAD, PISA SCORES ARE HIGHER.

Impact of teacher-directed instruction[1]

Average PISA science score with different amounts of teacher-directed instruction



1 Statistically significant in regression controlling student socioeconomic status, school type, and location. High point in regression is between most classes and all classes.

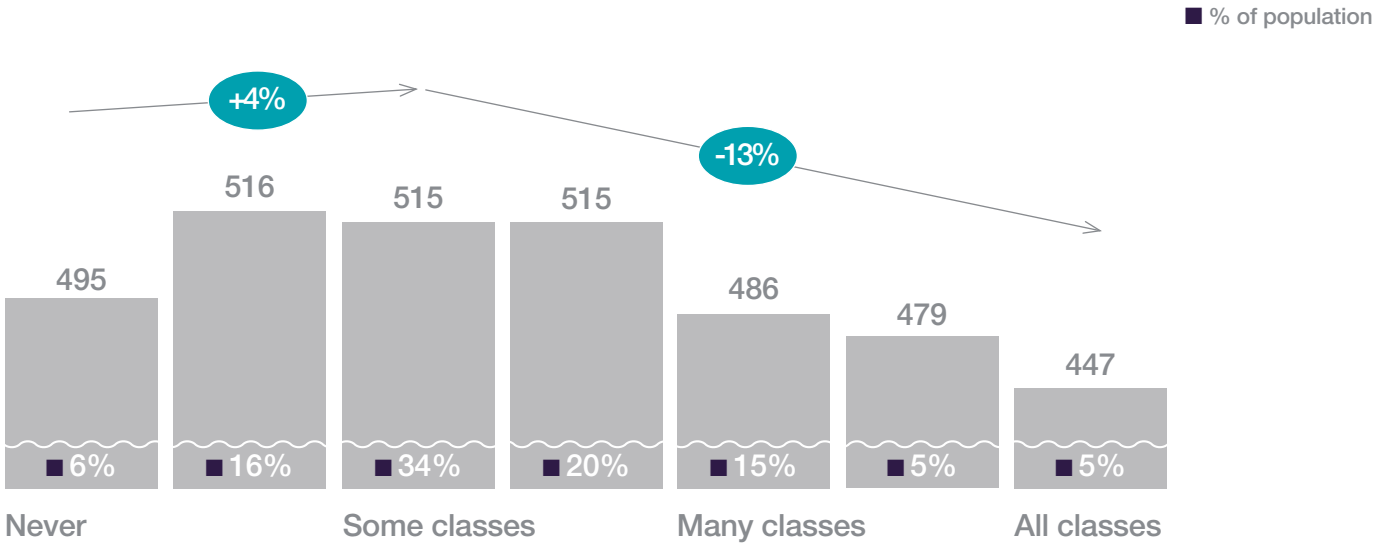
Source: OECD PISA 2015, McKinsey analysis

Based on PISA 2015 student survey responses in North America, scores rise with increased teacher-directed instruction. Students in classrooms where teacher-directed instruction is used "in most-to-all classes" score 14 percent higher than students in classrooms where teacher-direction is experienced "never or hardly ever" (Exhibit 7).

EXHIBIT 8: INQUIRY-BASED TEACHING DELIVERS MIXED RESULTS.

Impact of inquiry-based teaching[1]

Average PISA science score with different amounts of inquiry-based teaching



1 Statistically significant in regression controlling student socioeconomic status, school type, and location. High point in regression is around some classes, with rapid declines in score for many classes or more.
Source: OECD PISA 2015, McKinsey analysis

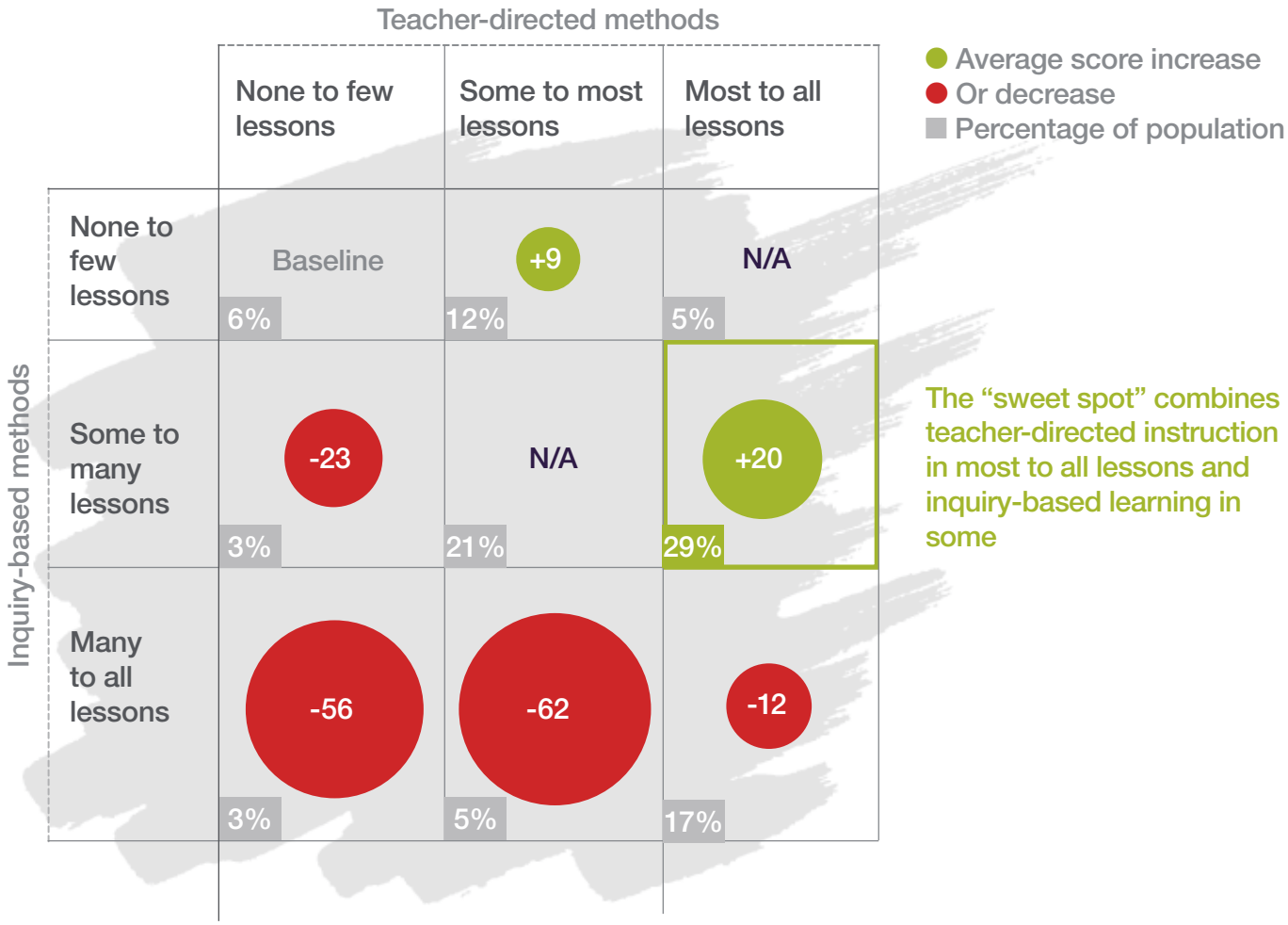
The picture for inquiry-based teaching is more subtle. While scores initially rise with some inquiry-based teaching, they then decrease with more frequent use (Exhibit 8).

At first blush, then, inquiry-based learning looks like a less effective choice. But when we dug into the data, we found a more interesting story: what matters is the interplay between the two types of teaching; there is a place for both. Inquiry-based learning can be effective—but only when strong teacher-directed teaching is in place, too. This suggests that teachers need to be able to explain scientific concepts clearly and students need to have content mastery in order to fully benefit from inquiry-based learning. Based on the 2015 PISA results, the most effective combination appears to be teacher-directed instruction in most or almost all classes, with inquiry-based learning in some-to-many. Students who receive this blend of teaching outperform those who experience high levels of inquiry-based learning without a strong foundation of teacher-directed instruction by 20 PISA points (Exhibit 9). To put it another way, the more teacher-directed instruction there is, the better it supports inquiry-based teaching.

This sweet spot is the same in every global region we studied. Also true in every region: excessive inquiry-based teaching (in many to all classes) without a strong foundation of teacher direction leads to the worst student outcomes. In North America, for example, students receiving too much inquiry-based teaching score 75–80 PISA points lower than the sweet spot. That is equivalent to two whole years of learning. What differed across regions was the degree to which the addition of inquiry-based teaching improved

EXHIBIT 9: FINDING THE SWEET SPOT: THE BEST STUDENT OUTCOMES OCCUR WHEN BOTH TEACHING STYLES ARE USED.

Impact of teacher-directed and inquiry-based combinations
Average point increase in North America PISA science score relative to baseline[1]



1 Expected change in score in a regression controlling for student socioeconomic status, school type and location. All results shown are statistically significant at 9% percent except for inquiry in many-to-all and teacher-directed in most-to-all (94% significance) and inquiry in none-to-few and teacher-directed in some-to-most (significant at 85% level). N/A denotes not significant at 85% level.
SOURCE: OECD PISA 2015, McKinsey analysis

scores over and above a purely teacher-directed approach. In developing regions such as Latin American and the Middle East and Africa, there was only a small additional benefit (1–2 PISA points) from adding inquiry-based teaching to a highly teacher-directed approach. In North America, the difference was 20 points.

Only 29 percent of North American students are sitting in classes with the optimal balance of teacher- and inquiry-based approaches. About half are receiving too little teacher-directed instruction, and roughly a quarter are receiving too much inquiry-based teaching. If all North American students moved

into the sweet spot, PISA scores would rise by 4.4 percent (22 PISA points, or more than a half a year of learning).

These results do not take into account how good the teaching itself is. There are certainly quality gaps in both teacher-directed and inquiry-based classrooms. It also bears remembering that there is a diverse range of inquiry-based practices, and these have measurably different effects (Exhibit 10).

The results suggest that school systems need to be careful in selecting inquiry-based teaching practices. More structured inquiry-based methods seem to have a generally positive impact on student outcomes. Every student should have the opportunity to conduct experiments and draw conclusions from those experiments. On the other hand, our analysis shows that some less structured techniques may actually hurt student outcomes in North America: among them, having students design their own experiments, having a class debate about investigations, and having students argue about science questions. These practices were associated with lower student scores, at any frequency, in every global region.

Given that there is strong support for inquiry-based pedagogy among education professionals, these findings may seem counter-intuitive. We offer two hypotheses for why inquiry-based teaching is not translating into better student outcomes. First, students cannot progress to inquiry-based methods without a strong foundation of knowledge gained through teacher-directed learning. Second, inquiry-based teaching might be more challenging to deliver effectively, and teachers who attempt it without sufficient training and support might struggle.

We should emphasize, too, that inquiry-based practices may bring benefits beyond improving student scores. The PISA data also shows that experiencing inquiry-based teaching increases students' joy in science and their belief that doing well in science will be worthwhile for their future (what PISA calls "instrumental motivation"). This matters because passion for a topic is linked to perseverance. Teacher-directed learning is also positively correlated with joy in science and instrumental motivation, but it does not have nearly as strong an impact.

Knowing all this is only the start, and raises many questions about how to find the right balance between teacher-directed and inquiry-based teaching, and how to improve the quality of each. At a minimum, our research suggests that teachers need to understand the content they are teaching, and be able to explain it, before they can jump into inquiry-based practices □

EXHIBIT 10: MORE STRUCTURED INQUIRY-BASED METHODS IMPROVE STUDENT OUTCOMES; LESS STRUCTURED METHODS APPEAR TO HURT THEM.

Impact of Inquiry-based practices

1 Regression output controlling for student socioeconomic status, school type, and location normalized over the intercept.
2 Result for this practice not statistically significant at 95% level for many classes.
SOURCE: OECD PISA 2015, McKinsey analysis

Finding 3:
Increasing the school day up to 6.5 hours can improve student outcomes, but significant gains can also be made from using existing time better

The average reported school day for 15-year-olds who took the 2015 PISA assessment in North America is 5.3 hours. Although the average school day is similar across the United States and Canada, there is more variation in the United States. In Canada, 79 percent of students are in class between five to six hours a day, in contrast to just 60 percent of US students. In the United States, 17 percent of students are in schools with fewer than five hours of instruction and 23 percent are in schools with more than six hours.

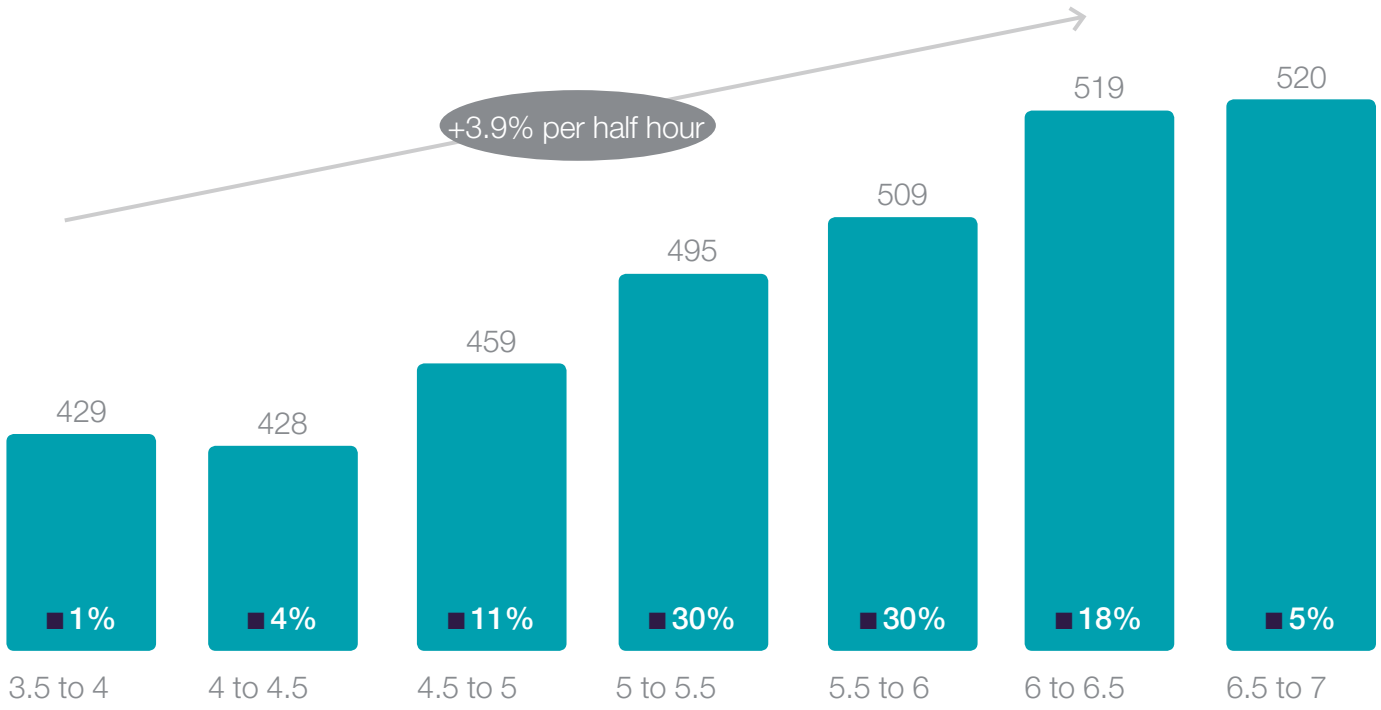
This diversity is no doubt partly a reflection of the highly decentralized US school system. It may also be a reflection of the ongoing push to increase learning time through extended school days and a longer academic year. The impetus for this has come from both the federal and state governments, supported by research suggesting that greater time in school is strongly associated with better student performance especially for at-risk students.[16] The variation in the length of the school day in the United States can help us to understand the implications of increasing hours of instruction. We looked to the PISA data to understand the academic impact of each half hour of additional instruction.

In the United States, PISA science scores increase 3.9 percent for each additional half hour of instruction, up to 6.5 hours per day. This relationship holds controlling for student socioeconomic status, school type, and location with the highest scores obtained by students with 6.0 to 6.5 hours of school per day (Exhibit 11).

If all students received similar outcomes to those currently receiving 6 to 6.5 hours of instruction per day, average PISA

EXHIBIT 11: EACH ADDITIONAL HALF HOUR OF INSTRUCTION IMPROVES STUDENT OUTCOMES.

PISA score benefit of additional instructional hours per day[1]
United States PISA science score 2015 by length of the school day



1 In a regression controlling for student socioeconomics, school type, and location the high point is at 6 to 6.5 hours per day.
Source: OECD PISA 2015, McKinsey analysis

science scores would increase by 4.4 percent (22 PISA points, equivalent to more than half a year of learning). However, from a policy perspective, increasing hours of instruction would be expensive and could negatively impact after-school activities, especially for children with long commutes. Extending the school day could also make it harder to maintain or shift to later start times, and there is a growing body of evidence to suggest that these improve graduation and attendance amongst high-school students.[17]

It's also worth noting that some European systems—such as in Finland and Germany—achieve superior results to the United States despite having a shorter average school day. This suggests that lengthening the school day is only part of the answer. Improving the quality of every hour in school also remains critical □

Finding 4:
Early childhood education had a positive academic impact on today's 15-year-olds in most regions of the world, but not in the United States

More than half of the synaptic connections that allow people to think, see, hear, and speak are formed before age three.[18] Although brain plasticity persists into adulthood, the brain is most receptive to interventions in early childhood. That is the promise of quality early childhood education (ECE), and research shows that such programs can indeed improve academic and social outcomes, especially for disadvantaged children.[19] High quality ECE programs can narrow the achievement gap by helping disadvantaged children gain cognitive, social, and other skills before starting kindergarten.

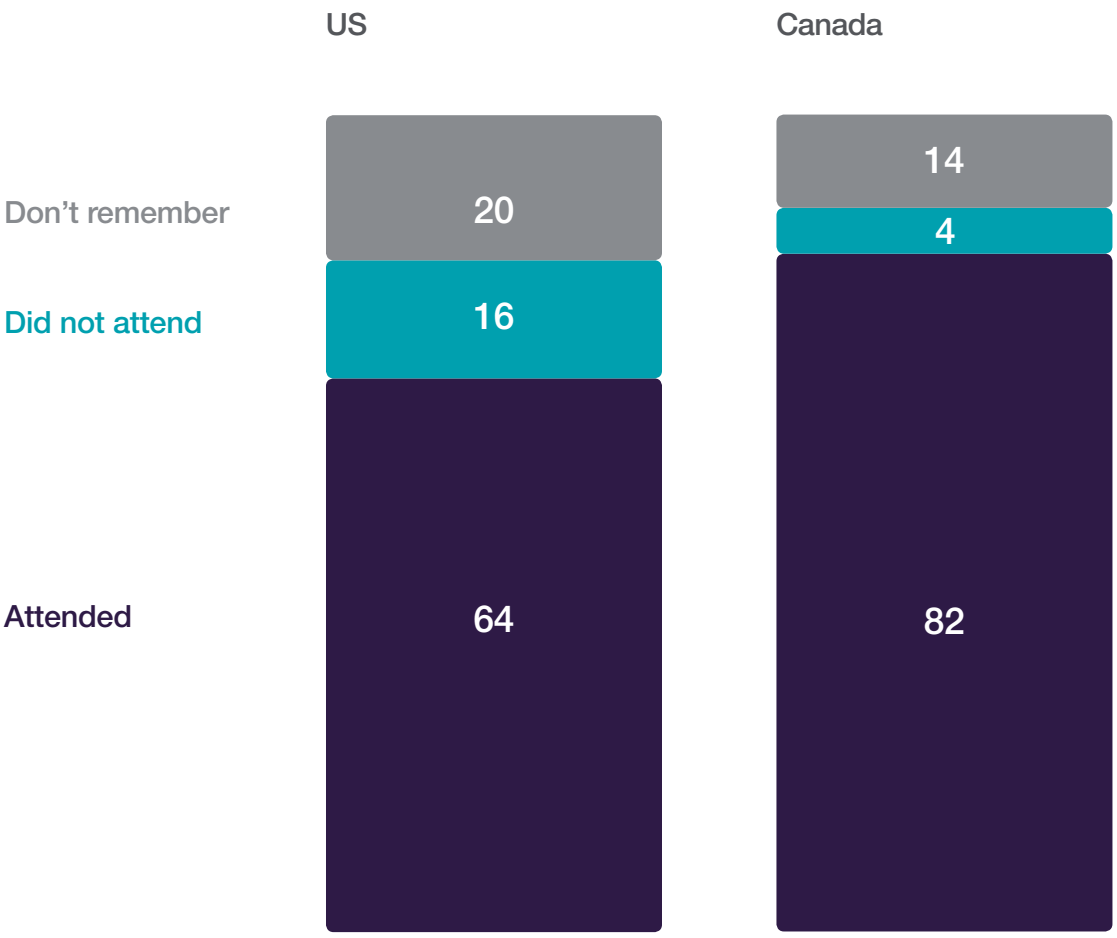
The PISA survey asked students how old they were when they started ECE. The answers to this question shed light both on the proportion of children attending ECE in North America, and the effect this participation has on academic outcomes at age 15.

In Canada, 82 percent of students report having attended ECE (starting at age five or younger), compared with 64 percent in the United States. This difference is echoed in UNESCO data, which also shows higher pre-primary enrollment in Canada (Exhibit 12).

EXHIBIT 12: AMERICAN CHILDREN ARE LESS LIKELY TO ATTEND ECE THAN CANADIANS.

Reported ECE attendance

Percent of 15-year-old students who report attending ECE[1]



1 In response to the question "At what age did you start <ISCED 0>" with "Did not attend" being the sum of not attending and starting at age 6+; "Attended" being the sum of starting at age 1 or younger, 2, 3, 4, or 5.This gap in ECE attendance is also reflected in UNESCO enrollment data which reports 64% pre-primary enrollment for US and 72% for Canada

Source: OECD PISA 2015; UNESCO; McKinsey analysis

EXHIBIT 13: IN THE UNITED STATES, CHILDREN WHO ATTENDED ECE HAVE LOWER PISA SCORES; IN OTHER REGIONS, THEY SCORE BETTER..

Impact of early childhood education
PISA point difference between children with and without ECE[1]

1 Controlling for student socioeconomic status, school-type, and location.
2 Average of EU and non-EU.
Source: OECD PISA 2015, McKinsey analysis

In most regions of the world, attending ECE is associated with a small but statistically significant increase in PISA science score at age 15, even after controlling for student socioeconomic status, school type, and location. This is also the case in Canada, where 15-year-olds who had attended ECE score 3 percent higher (or 15 PISA points) than those who did not (Exhibit 13). In the United States, however, students who attended ECE programs actually score 2 percent lower (8 PISA points).

Our hypothesis is that this boils down to a quality problem. Research shows that although high quality programs boost outcomes, low quality early childhood programs are less effective, with any benefits more likely to fade out over time.[20]

EXHIBIT 14: ECE IS FALLING SHORT FOR LOWER SOCIOECONOMIC STATUS STUDENTS IN THE UNITED STATES.

	Impact of early childhood Increase in PISA science score[1]		**Students with no ECE** % of students	
	US	Canada	US	Canada
High socioeconomic status	+2%	+3%	9%	4%
Upper-middle socioeconomic status	-2%	+4%	14%	3%
Lower-middle socioeconomic status	-3%	+3%	18%	4%
Low socioeconomic status	-2%	+3%	20%	6%

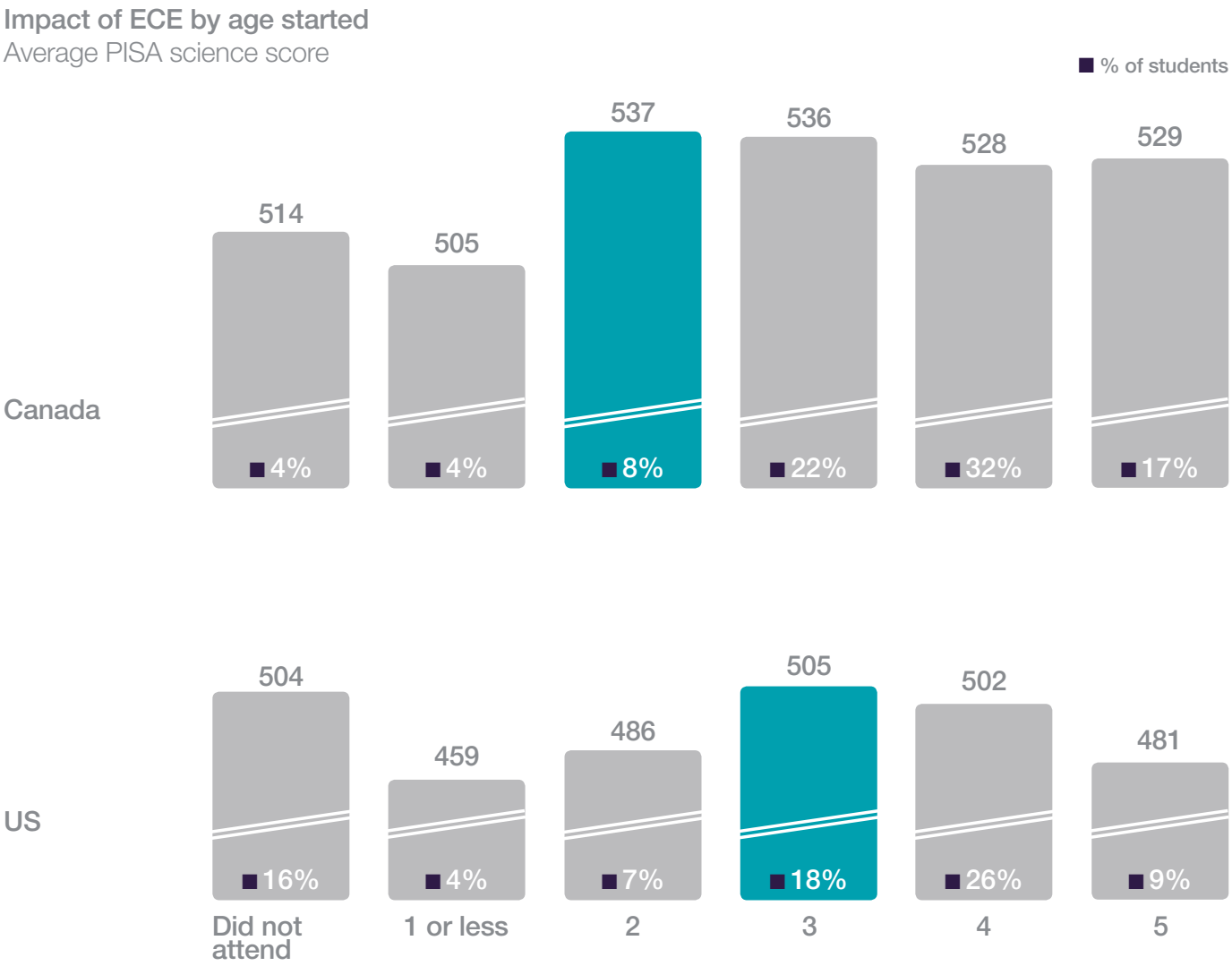
1 Descriptive % difference in score; In a regression controlling for student socioeconomic status, school type and location US students from ESCS 1+2 (bottom half socioeconomic status) score 11 PISA points lower if they attend ECE. (95% confidence). US students from ESCS 3+4 (top half) score 4 PISA points higher if they attend ECE (The difference in ECE impact of 15 PISA points is significant at 94% confidence interval). In Canada there is no statistically significant difference between the socioeconomic quartiles in impact from attending ECE

Source: OECD PISA 2015, McKinsey analysis

Digging deeper into the data, we can test this by examining the impact of ECE by socioeconomic quartile. In the United States, children from lower socioeconomic quartiles are much less likely to attend ECE, and those that do attend do not see academic benefits, whereas higher socioeconomic status students do benefit. In Canada, students from all socioeconomic backgrounds attend at similar rates, and benefits are also similar at all levels (Exhibit 14).

Based on our research, ECE for lower socioeconomic status children is falling short in the United States. Many studies have shown that, although poor children in the United States benefit from high-quality ECE,[21] they are less likely to attend the highest-quality programs.[22] Lower-middle-income children appear to be benefiting least from ECE. This may be because their families are too wealthy to qualify for government programs, but too poor to afford high-quality private ones.

EXHIBIT 15: BEST AGE TO START ECE IS YOUNGER IN CANADA, BUT STARTING BEFORE AGE TWO IS HARMFUL IN BOTH COUNTRIES



1 Regression results controlling for socioeconomic status, school type, and location; co-efficient plus intercept normalized to average country score.
2 Numbers do not add up to 100% as analysis excludes students who "did not remember" (14% in Canada and 20% in the US).
SOURCE: OECD PISA 2015

The age of starting ECE also matters. Students in Canada do best when they start at age 2 whereas students in the United States do best when they start at age 3 (Exhibit 15). Starting too young is problematic in both countries, but especially so in the United States. In fact, students who start ECE at age 1 or younger in the United States score 45 PISA points lower at age 15 than students who do not attend at all. Students in Canada starting ECE at age 1 or less score 9 PISA points less than those who do not attend.

Of course, ECE is about more than scoring well on PISA tests many years later. Non-cognitive attributes such as physical and mental health, attainment metrics including graduation rates, as well as social and emotional development, are also important. Also, more than a decade has passed since these students were in ECE, and several American states have made significant investments since then.

Nonetheless, our analysis of the PISA data suggests that there may be a significant ECE quality problem in the United States, especially for low-income children and very young children. As systems consider enrolling more and younger children, they should consider what alternatives this care is replacing, as well as carefully monitoring the quality of both public and private ECE. □

Conclusion

Our research has for the first time applied machine learning and advanced analytics to the OECD PISA dataset to understand the drivers of student outcomes in North America. This analysis has shed important light on some critical and oft-debated areas of education, and also identified new territories worthy of further exploration.

For each of the four findings, there is a clear need for additional research. Within mindsets, the priority is to determine what system-level interventions can make a difference in shifting student mindsets, and what effect these interventions have on student outcomes.

For teaching practices, more research is needed into how to effectively combine teacher-directed and inquiry-based learning. Across the board, more research is needed on how to strike the right balance between increasing access and improving quality—particularly relevant in terms of

increasing hours of instruction and extending access to early childhood education. In an important sense, then, this report is part of a longer journey.

With its emphasis on data and analysis, this research aims to help North American school systems move from good to great and beyond. Even a survey as large and rigorous as the PISA data set provides only some of the answers. But we believe that the four findings outlined here, combined with the conclusions of our 2010 report on the world's most improved school systems, provide useful insights to guide North American policymakers as they continue the work of improving the education and thus the lives of the region's students ■

analytical appendix

To analyze the PISA dataset, we used a variety of modern machine learning and traditional statistical techniques.

First, we used SparkBeyond, an automated feature-discovery engine that uses large-scale combinatorial testing of millions of transformations on raw data to identify relevant drivers of outcomes—in our case, PISA student scores. SparkBeyond can create features from numeric, time series, text, and other inputs, and works best with complex data sets with thousands of variables and millions of data points. For the 2015 OECD PISA data, this entailed testing more than 1,000 survey variables derived from student, teacher, parent, and principal surveys for the approximately 540,000 students who took the PISA examination. This identified variables and groups of variables that were most predictive of student performance.

We excluded from our SparkBeyond and subsequent analysis highly predictive variables where the direction of causality was strongly in question, including grade repetition, student self-efficacy, environmental awareness, expected educational attainment, and epistemological beliefs.

We then carried out traditional descriptive and predictive statistical analyses on the identified features that were most important in determining performance both within 2015 dataset and across the PISA surveys since 2000.

For every analysis, we tested whether findings held in a regression controlling for economic, social, and cultural status (ESCS), type of school (SC013Q01: is your school a public or private school school?) and location of school (SC001Q01: which of the following definitions best describes the community in which your school is located?).

For the 2015 OECD PISA data, this entailed testing more than 1,000 survey variables derived from student, teacher, parent, and principal surveys for the approximately 540,000 students who took the PISA examination.

Where the regression results were consistent with the descriptive analysis, we have used the descriptive analysis in the report. Where the regression tells a different story from the description, we have reported regression coefficients to preserve the rigor of our findings.

We also tested our insights by school and student segment, creating two more screens—specifically, school performance level and student socioeconomic status.

School performance:

we used the numerical cut-offs from our 2010 report to define poor, fair, good, great, and excellent school systems. Each category represents approximately one school-year equivalent, or 40 PISA points.

- Excellent: >560 points
- Great: 520-560 points
- Good: 480-520 points
- Fair: 440-480 points
- Poor: <440 points

Distribution of students by school performance level

Students	Poor (%)	Fair (%)	Good (%)	Great (%)	Excellent (%)
N America	14	23	39	18	5
Latin America	76	15	6	2	0
Non-EU	35	21	29	13	3
EU	18	20	28	20	14
MENA	89	8	2	1	0
Asia	43	16	15	13	13

Then we applied these cut-offs to individual schools as well as to school systems. We did this because there may be pockets of poorly performing schools in otherwise good systems. In these schools, the interventions applicable to poor systems may apply, even if they are in a country that on the whole performs at a "good" level. Based on this analysis, we could determine the percentage of students in differently performing schools for each region and country

Student socioeconomic status:

We use the term "student-socioeconomic-status quartile" throughout the report. This refers to PISA's ESCS indicator that integrates a number of measures related to students' backgrounds, including their parents' occupations, education levels, and possessions. We created ESCS quartiles by region based upon student weights.

Target variables and plausible values

We used the 2015 PISA science score as the target variable because the 2015 test focused on science both for the assessment and survey questions (in 2012, PISA focused on math, and in 2009, on reading). To calculate the PISA science score at the student level, we averaged the results of all the plausible values for science (PV1 to PV10 for science).

To roll up scores at the regional level, we used student weights to represent each country based on its student population. For example, the Latin American numbers all refer to weighted average student scores across Latin America; the same is true for all other regions.

For consistency with OECD publications, we used a slightly different methodology in the overview of historical regional performance. This approach creates a country-level average, first using student weights (such as "average score for Brazil"), but then takes the straight average of the scores of countries in a particular region or a group (such as "all OECD countries").

Description of specific variables

In addition to using existing OECD PISA variables and indices, we created our own indices for some analyses.

Motivation calibration:

Motivation calibration is a measure of a student's ability to recognize motivation in others, or the extent to which the student's definition of motivation agrees with the standard definition. Specifically, we took the PISA question ST121, which presented three student archetypes and asked the respondent to what extent they agree that each archetype is motivated on a four-point scale, ranging from "strongly disagree" to "strongly agree."

Based on our assessment of the motivation level of each archetype, we assigned a weight of -2 to the first student (NAME 1—highly unmotivated), +1 to the second student (NAME 2—somewhat motivated), and +2 to the third student (NAME 3—highly motivated).

For example, a student who strongly disagreed that <NAME 1> is motivated, agreed that <NAME 2> is motivated, and strongly agreed that <NAME 3> is motivated would accumulate the following score:

- 1 * -2 = -2: one point for strongly disagree with a weight of -2 for <NAME 1>
- 3 * 1 = 3: three points for agree with a weight of 1 for <NAME 2>
- 4 * 2 = 8: four points for strongly agree with a weight of 2 for <NAME 3>
- Total score: -2 + 3 + 8 = 9

ST121 **Please read the descriptions about the following three students. Based on the information provided here, how much would you disagree or agree with the statement that this student is motivated?** (Please select one response in each row.)

		Strongly disagree	Disagree	Strongly Agree	Agree
ST121Q01NA	<NAME 1> gives up easily when confronted with a problem and is often not prepared for his classes. <Name 1> is motivated.	☐1	☐2	☐3	☐4
ST121Q02NA	<NAME 2> mostly remains interested in the tasks she starts and sometimes does more than what is expected from her. <Name 2> is motivated.	☐1	☐2	☐3	☐4
ST121Q03NA	<NAME 3> wants to get top grades at school and continues working on tasks until everything is perfect. <Name 3> is motivated.	☐1	☐2	☐3	☐4

ST034 **Thinking about your school: to what extent do you agree with the following statements?** (Please select one response in each row.)

		Strongly disagree	Disagree	Strongly Agree	Agree
ST034Q01TA	I feel like an outsider (or left out of things) at school.	☐1	☐2	☐3	☐4
ST034Q02TA	I make friends easily at school.	☐1	☐2	☐3	☐4
ST034Q03TA	I feel like I belong at school.	☐1	☐2	☐3	☐4
ST034Q04TA	I feel awkward and out of place in my school.	☐1	☐2	☐3	☐4
ST034Q05TA	Other students seem to like me.	☐1	☐2	☐3	☐4
ST034Q06TA	I feel lonely at school.	☐1	☐2	☐3	☐4

ST119 **To what extent do you disagree or agree with the following statements about yourself?** (Please select one response in each row.)

		Strongly disagree	Disagree	Strongly Agree	Agree
ST119Q01NA	I want top grades in most or all of my courses.	☐1	☐2	☐3	☐4
ST119Q02NA	I want to be able to select from among the best opportunities available when I graduate.	☐1	☐2	☐3	☐4
ST119Q03NA	I want to be the best, whatever I do.	☐1	☐2	☐3	☐4
ST119Q04NA	I see myself as an ambitious person.	☐1	☐2	☐3	☐4
ST119Q05NA	I want to be one of the best students in my class.	☐1	☐2	☐3	☐4

We defined a cutoff of 8 points in the aggregated score, which ensures that only the following students are classified as having a strong motivation calibration:

- Students who strongly agree that <NAME 3> is motivated, and whose agreement on <NAME 1>'s motivation does not exceed their agreement on <NAME 2>'s motivation
- –OR— Students who agree that <NAME 3>'s is motivated; agree that <NAME 2> is motivated, and strongly disagree that <NAME 1> is motivated
- –OR— Students who agree that <NAME 3> is motivated; strongly agree that <NAME 2> is motivated, and disagree or strongly disagree that <Name 1> is motivated

Sense of belonging:

We grouped the index BELONG (based on ST034) as follows:

- Low belonging: BELONG < 0
- High belonging: BELONG >=0

Motivation:

We grouped the index MOTIVAT (based on ST119) as follows:

- Low belonging: MOTIVAT < 0
- High belonging: MOTIVAT >=0

Test anxiety:

We grouped the index ANXTEST (based on ST118) as follows:

- Low belonging: ANXTEST < 0
- High belonging: ANXTEST >=0

Instrumental motivation:

We grouped the index INSTSCIE (based on ST113) as follows:

- **Low instrumental motivation: INSTSCIE < 0**
- **High instrumental motivation: INSTSCIE >=0**

Growth vs. fixed mindset:

To assess the impact of a growth versus fixed mindset, we used selected 2012 PISA survey question ST43 and ST91 from the student survey.

We created an index by adding the response values for each of the four sub-questions related to growth versus fixed mindsets, after reversing the sequence of response values for the last question to account for the negative framing of the prompt.

The resulting index takes values from 4 to 16, with lower scores representing a growth mindset and higher scores representing a fixed mindset. Looking at the distribution of students globally, we devised the following definitions.

- **Strong growth mindset: students with a score of 4 or 5 reflect a growth mindset on at least three of the sub-questions, and are directionally aligned on the remaining question. These represent 23 percent of the global population.**

- **Neutral or weak growth mindset: students with a score of 6 to 9 reflect a neutral or weak growth mindset and represent 69 percent of the global population.**

ST118 **To what extent do you disagree or agree with the following statements about yourself?** (Please select one response in each row.)

		Strongly disagree	Disagree	Strongly Agree	Agree
ST118Q01NA	I often worry that it will be difficult for me taking a test.	☐1	☐2	☐3	☐4
ST118Q02NA	I worry that I will get poor <grades> at school.	☐1	☐2	☐3	☐4
ST118Q03NA	Even if I am well prepared for a test I feel very anxious.	☐1	☐2	☐3	☐4
ST118Q04NA	I get very tense when I study for a test.	☐1	☐2	☐3	☐4
ST118Q05NA	I get nervous when I don't know how to solve a task at school.	☐1	☐2	☐3	☐4

ST113 **How much do you agree with the statements below?** (Please select one response in each row.)

		Strongly disagree	Disagree	Strongly Agree	Agree
ST113Q01TA	Making an effort in my <school science> subject(s) is worth it because this will help me in the work I want to do later on.	☐1	☐2	☐3	☐4
ST113Q02TA	What I learn in my <school science> subject(s) is important for me because I need this for what I want to do later on.	☐1	☐2	☐3	☐4
ST113Q03TA	Studying my <school science> subject(s) is worthwhile for me because what I learn will improve my career prospects.	☐1	☐2	☐3	☐4
ST113Q04TA	Many things I learn in my <school science> subject(s) will help me to get a job.	☐1	☐2	☐3	☐4

ST043 **Thinking about your math lessons: to what extent do you agree with the following statements?** (Please select one response in each row.)

		Strongly disagree	Disagree	Strongly Agree	Agree
(a)	If I put in enough effort I can succeed in mathematics	☐1	☐2	☐3	☐4
(b)	Whether or not I do well in maths is up to me	☐1	☐2	☐3	☐4
(c)	If I wanted to, I could do well in mathematics	☐1	☐2	☐3	☐4
(d)	I do badly in mathematics whether or not I study for my exams	☐1	☐2	☐3	☐4

SC004 **The goal of the following set of questions is to gather information about the student-computer ratio for students in the <national modal grade for 15-year-olds> at your school.**
(Please enter a number for each response. Enter "0" [zero] if there are none.)

		Number
SC004Q01TA	At your school, what is the total number of students in the <national modal grade for 15-year-olds>?	
SC004Q02TA	Approximately, how many computers are available for these students for educational purposes?	
SC004Q03TA	Approximately, how many of these computers are connected to the Internet/ World Wide Web?	
SC004Q04NA	Approximately, how many of these computers are portable (e.g. laptop, tablet)?	
SC004Q05NA	Approximately how many interactive whiteboards are available in the school altogether?	
SC004Q06NA	Approximately how many data projectors are available in the school altogether?	
SC004Q07NA	Approximately how many computers with internet connection are available for teachers in your school?	

ST125 **How old were you when you started <ISCED 0>?**
(Please choose from the drop-down menu to answer the question.)

Years

Please choose ▼
Option A
Option B
Option C
Option …

Drop-down menu, offering answers "1 year or younger", 2 years, 3 years, 4 years, 5 years, "6 years or older", "I did not attend <ISCED 0>", "I do not remember".

- **Fixed mindset: students with a score of 10 to 16 have an average response of 2.5 or more on the four questions, meaning that they tend to be misaligned with the principles of a growth mindset. They represent 8 percent of the global population.**

We compared students with a fixed mindset to students with a strong growth mindset in our analysis. In addition, we found that incremental gains were seen at each stage from fixed to neutral and from weak growth to strong growth.

Teaching practices:

To assess teaching practices, the PISA survey asked a series of questions about teacher-directed instruction (ST103) and inquiry-based instruction (ST098). This question does not allow us to assess the intensity of the teaching practices in a given class, but only the frequency with which they occur.

Students responded on a frequency scale that was slightly different for each set of questions:

Teacher-directed learning (ST103)
1 = Never or almost never
2 = Some lessons
3 = Many lessons
4 = Every lesson or almost every lesson

Inquiry-based learning (ST098)
1 = In all lessons
2 = In most lessons
3 = In some lessons
4 = Never or hardly ever

We consolidated each student's responses into averages on a scale from 1 to 4—one average for teacher-directed instruction and another for inquiry-based instruction (with the numbers reversed to be comparable). These averages form the basis for our analysis of teaching practices.

The OECD also created a numerical index of teacher-directed (TDTEACH) and inquiry-based learning (IBTEACH), which is calibrated such that the OECD average is 0 and the standard deviation is 1. When we ran regressions on the TDTEACH and IBTEACH variables, our results were consistent with theirs. However, we chose to present the data using our own indices because we believed these gave a clearer picture what was happening in the classroom.

ICT at school:

to create a like-for-like comparison of the impact of ICT hardware, we used the survey questions asked of school principals from SC004 and normalized the results by classroom size and student-to-teacher ratio. This allowed us to evaluate the effect adding one projector, student computer, or teacher computer to an average class size of 36 students.

Early childhood:

To understand the impact of early-childhood education (ECE) we used the student survey question ST125. We excluded from the analysis students who could not remember when they started ECE. With the remaining students, we counted them as having attended ECE if they started at five years or younger. Students who started at six years or older or who responded "no early-childhood education" we counted as not having attended ECE. Note we did not use the simpler question ST124 ("Did you attend early-childhood education," as only 15 percent of students globally answered this question (versus 82 percent who answered ST125). We also cross-checked results against similar questions in the parent survey for the subset of countries that took the parent survey; the results were consistent ■

1 These five stages inform McKinsey's Universal Scale of education system performance, which combines available assessments like PISA, TIMSS, TERCE, or local tests onto a single scale. We normalized the data, creating new units that are equivalent to 2000 PISA scores, and then broke down the results into five categories: poor, fair, good, great and excellent (see the analytical appendix for more, and also http://www.mckinsey.com/industries/social-sector/our-insights/how-the-worlds-most-improved-school-systems-keep-getting-better).

2 Argentina, Kazakhstan, and Malaysia were excluded from PISA 2015 report, but are included in our analyses. The PISA 2015 sample for Malaysia did not meet PISA response-rate standards; the PISA 2015 sample from Argentina did not cover the full target population; and the results from Kazakhstan are based only on multiple-choice items. Because our report analyzes achievement drivers at the student level based on examining individual items, and reports are generated at a regional basis rather than comparing the performance of individual countries, we have included these countries in our analysis. Albania was excluded from our analysis because, due to the ways in which the data were captured, it was not possible to match the data in the test with the data from the student questionnaire. As our report is entirely based on the drivers from the student questionnaire, we could not include Albania in our analysis.

3 The 2015 computer-based assessment was designed as a two-hour test comprising four 30-minute clusters. Students took two science clusters, plus two others across reading, math, and collaborative problem solving.

4 OECD PISA Country Note: Key Findings from PISA 2015 for the United States.

5 Claro, S., Paunesku, D., & Dweck, C. S., "Growth mindset tempers the effects of poverty on academic achievement," Proceedings of the National Academy of Sciences, 113(31), 8664-8668, doi:10.1073/pnas.1608207113.

6 Duckworth, Angela, Grit: The Power of Passion and Perseverance, New York, NY: Scribner, 2016.

7 Credé, M., Tynan, M. C., & Harms, P. D., "Much ado about grit: A meta-analytic synthesis of the grit literature," Journal of Personality and Social Psychology, 2016, doi:10.1037/pspp0000102; and The impact of non-cognitive skills on outcomes for young people, Education Endownment Foundation, 2013, https://educationendowmentfoundation.org.uk/public/files/Publications/EEF_Lit_Review_Non-CognitiveSkills.pdf.

8 To attain statistically meaningful results, we selected the top 100 variables using a feature-identification machine-learning algorithm. Recognizing that the regression wouldn't distinguish collinearity across variables, we mitigated by placing variables very likely to be collinear in the same category. We cannot control for collinearity between categories.

9 Asia, Europe, Middle East & North Africa, Latin America, and North America

10 Tang, Sandra et al, "Adolescent pregnancy's intergenerational effects: Does an adolescent mother's education have consequences for her children's achievement?" Journal of Research on Adolescence, 2014; Harding, Jessica et al, "The relationship between maternal education and children's academic outcomes: A theoretical framework," Journal of Marriage and Family, 2015.

11 Gordon, Robert, Thomas J. Kane, and Douglas O. Staiger, Identifying Effective Teachers Using Performance on the Job, The Hamilton Project White Paper, 2006; Hanushek, Eric, "Valuing teachings: How much is a good teacher worth?" Education Next, 2011; Chetty, Raj et al, "Measuring the Impact of Teachers," American Economic Review, 2014.

12 Paunesku, David, et al, "Mindset interventions are a scalable treatment for academic underachievement," Psychological Science, 1–10.

13 "Meta-cognition and self-regulation," Teaching & Learning Toolkit, Education Endowment Foundation, November 2017,

14 Excellence and equity in education, volume 1, OECD PISA, 2015.

15 A student's achievement at age 15 results from the accumulated skill and practice of the teachers she has had up to that point in her educational career, and the practice of a given teacher in that particular year's impact is additive to that accumulated experience. Nevertheless, we find strong relationships between PISA achievement and teacher practice in the PISA year.

16 Fryer, Roland and Dobbie, Will, "Getting beneath the veil of effective schools: Evidence from New York City," NBER Working Paper, 2011; Patall, Erika, Cooper, Harris, and Allen, Asheley Batts, "Extending the school day or school year," Review of Educational Research, 2010; Farbman, David, "The case for improving and expanding time in school," National Center on Time & Learning, 2015.

17 McKeever, Pamela Malaspina, "Delayed high school start times later than 8.30am and impact on graduation rates and attendance rates," Sleep Health Journal, 2017.

18 Shonkoff, Jack P. and Deborah A. Phillips, eds., From Neurons to Neighborhoods: the Science of Early Development, Board on Children, Youth, and Families: National Research Council and Institute of Medicine, Washington D.C.: National Academy Press, 2000.

19 Carneiro and Heckman, Minnesota Federal Reserve Bank, "The Region," December 2003; Arthur Reynolds et al., "Age 21 Cost Benefit analysis of the Title 'Chicago Child-Parent Center Programs," Waisman Center—University of Wisconsin Madison, June 2001; Mann, E., Reynolds, A., Robertson, D., and Temple, J., "Age 26 Cost-Benefit Analysis of the Title I Chicago Child-Parent Center Program;" Lifetime Effects: The High/Scope Perry Preschool Study Through Age 40," High/Scope, 2005; The Carolina Abecedarian Project: website; Carneiro and Heckman 2003; Center for Public Education website; HeadStart fade out:

20 Lowe Vandell, Deborah et al (2010) "Do Effects of Early Child Care Extend to Age 15 Years? Results from the NICHD Study of Early Child Care and Youth Development", Child Development, 81(3): 737–756. and Mashburn AJ, et al. (2008) Measures of classroom quality in prekindergarten and children's development of academic, language, and social skills. Child Development, 79: 732-749.

21 Gormley, W. T., Gayer, T., Phillips, D., & Dawson, B., "The effects of universal pre-k on cognitive development," Developmental Psychology, 41, 872–884, (2005); Magnuson, K., Ruhm, C., & Waldfogel, J., "The persistence of preschool effects: Do subsequent classroom experiences matter?" Early Childhood Research Quarterly, 22, 18–38, 2007; Weiland, C., & Yoshikawa, H., "Impacts of a prekindergarten program on children's mathematics, language, literacy, executive function, and emotional skills," Child Development, 84, 2112–2130, 2013.

22 Barnett et al, Equity and excellence: African-American children's access to quality preschool, Washington, D.C.: Center for Enhancing Early Learning Outcomes & National Institute for Early Education Research, 2013; Early, D. M. et al, "How do pre-kindergarteners spend their time? Gender, ethnicity, and income as predictors of experiences in pre-kindergarten classrooms," Early Childhood Research Quarterly, 25, 177–193, 2010; Bassok, D., Fitzpatrick, M., Greenberg, E., & Loeb, S., "The extent of within and between-sector quality differences in early childhood education and care", Unpublished manuscript, 2013.

Other studies in this series include: Asia | Europe Middle East & North Africa & Latin America