Introduction

By Professor Ray Fisman

The conversation about how to improve American education has often taken on a decidedly confrontational tone. The caricature presented in the press depicts data-minded reformers who see the solution as getting rid of low-performing teachers, standing off against unions who are opposed to almost any teaching metric and care more about their jobs than the children they’re supposed to be educating. Presenting the job of a teacher in such extreme terms does little to foster a constructive conversation about improving the quality of American education and fails to recognize the critical importance of developing classroom-tested, replicable and scalable approaches to teacher training that would make the educators we already have perform better. The goal of this half-day discussion at Columbia University’s Faculty House, presented in collaboration with the Robin Hood Foundation, was to expose leading education practitioners to new and innovative ways of improving teacher quality.

We don’t yet have all the answers, but the morning’s discussion provided insights from the many promising avenues that the presenters and their organizations are pursuing, and we hope allowed participants to take a step back and see their own organizations’ activities in a new light.

How might frontier research in social psychology inform the way teachers interact with their students? What type of teacher feedback is most effective? And what do we really know about what makes for effective teaching? If we ended up with more questions than answers, that was part of the point as well.

By bringing together a diverse group of education-focused leaders from the New York area, we hope ultimately to foster the further exchange of ideas and collaboration with the goal of improving education for students in New York and beyond.

Featured Speakers

In order of agenda

Ray Fisman
Lambert Family Professor of Social Enterprise
Co-director, Social Enterprise Program
Columbia Business School

Valerie Purdie-Vaughns
Assistant Professor of Social Psychology
Columbia University
Director, Intergroup Relations and Diversity Laboratory (IRDL)

John King
Commissioner of Education
President, University of the State of New York

Thomas Kane
Professor of Education and Economics
Harvard Graduate School of Education
Faculty Director, Harvard’s Center for Education Policy Research

Emary Aronson
Managing Director of Education
Robin Hood Foundation

Written by Jennifer Itzenson
Photography by Leslye Smith
Design by Heather Barbakoff
The latest statistics on public schools are not encouraging. More than a quarter of New York State’s high school students fail to graduate in four years, according to the State Education Department. More tellingly, however, only 25 percent of students graduate high school in four years and 48 percent of those graduates are college- and career-ready, among African-American students, those percentages are 56 percent and 15 percent.

The need to improve the quality of public school education—and the outcome for students—could not be more urgent. In recent decades, the labor market has become more demanding and the majority of jobs require at least some qualification beyond a high school degree. With harmful economic eroding opportunity, finding a job that provides a living wage, benefits, and opportunities for advancement is extraordinarily difficult (Figure 1).

While there is widespread agreement about the need to improve the education system, there is little consensus on the best way to carry out reforms—or even what those reforms should be. Some critics say the solution is to shut down failing schools, or at least fire the worst performing teachers. However, for many cities and towns, such a move would be politically impossible. For these places and for the nation as a whole the more meaningful challenge lies not in the firing and hiring of teachers, but in taking a given set of teachers and giving them the resources and support they need to become successful.

In his keynote presentation, John King, New York State’s commissioner of education, discussed the state’s ambitious reform agenda. The state is focused on increasing the graduation rate, and on ensuring students graduate prepared for college and careers. Currently, more than 50 percent of students in New York State’s four-year institutions of higher education are forced to take at least one remedial course. King noted, “An epidemic of remediation is a euphemism: it means that students are taking a high school class, for which they are paying college prices,” he said. Remedial classes often mean remediation, finding a job that provides a living wage, benefits, and opportunities for advancement is extraordinarily difficult (Figure 1).

Under King, New York State has undertaken a number of initiatives aimed at improving the quality of classroom instruction and reducing teacher turnover. These include the adoption of the Common Core, a set of standards that identifies the skills students need to acquire from the earliest grades through high school in order to be on track not only for remediation toward college and career success. The state will also expand the state’s data systems that allow teachers and principals to assess educational practice in real time. The reform agenda also includes a commitment to retrain, recredential and reward effective teachers and principals, and to improve the lowest-performing schools.

Differing from its predecessor, the Common Core, which is now in effect in 46 states, has also been committed to significant shifts in public school instruction. In reading, students now study a balance of informational and literary texts and are spending more time learning how to use evidence from source materials. In math, there is a new focus on coherence, fluency and application. These shifts are motivated by the skills needed for students to succeed in higher education and in the workforce. For a lot of people now, the United States has led the world in patterns, King said. But “to continue as we are, we have to hit a trail in fractions, and it turns out that fractions matter a lot more for later success in mathematics than patterns.”

Concurrent changes have emphasized improving the quality of classroom delivery. Too often, assessment has been limited to a “postmortem” analysis of a teacher’s effectiveness. King said. Data-driven instruction is a means of improving teacher effectiveness and student performance throughout the school year. As research by Jonah Rockoff, Associate Professor of finance and economics at Columbia Business School, and others have shown, a teacher’s effectiveness can have a profound impact on student achievement. Students taught by teachers who are in a series of highly effective teachers will do better, and students who have a series of ineffective teachers will do worse, King said. “And that’s why our assessments have become more challenging. The reality is that people do what is measured, and we must measure what we value.”

If meaningful performance feedback is a necessary input into helping teachers get better, evaluations must convey meaningful content, which has not traditionally been the case. Before the recent changes in the laws governing teacher evaluations, close to 99 percent of teachers were rated as at least satisfactory, according to the nonprofit The New Teacher Project. At the same time, most teachers agree that there is at least one poor performing teacher at their school. Improving the evaluation system will involve some sweeping changes, such as multiple observations by multiple observers, King said. Yet given the promising results from research on the improvements that come from evaluation-driven feedback, the benefits are worth the effort.

The state’s reform agenda is ambitious, King conceded. He is often urged to implement changes more slowly—adopt the Common Core, perhaps, but put off a new evaluation system. However, students cannot afford further delays, King argued. And the changes reinforce one another, so there is value to their simultaneous implementation. “Students are accountable for the Common Core today,” he said. “When they go to college and they’re sent to a remedial course, or they can’t get a job as a welder because they don’t understand fractions, they’re accountable for the Common Core. Every student deserves the best possible education, and our job is to deliver that.”

In the question-and-answer session that followed King’s presentation, he was asked about plans to link teacher performance to compensation. While the state has launched pilot programs to reward effective teachers, this remains a difficult area for reform. King said. “Given the history of the labor movement, with its emphasis on solidarity, it is a cultural challenge to talk about differentiating rewards,” he said. “Over time, I think that our compensation model will have to change if we want to attract the strongest possible teachers.” The state’s research has shown that it is important to link working conditions and opportunities for growth even more than compensation, he added. “Creating a performance management system in schools will lead to better working conditions,” he said. “And that will not only lead people to become teachers, but to stay teachers.”

Ray Fisman, co-director of the Social Enterprise Program, asked about efforts that were tried but found to be less than successful. King discussed the struggles some districts are facing in managing the time necessary for evaluations. “Many people think that the way to do an observation is to have an hour-long conference, and then observe for an hour, and then have an hour-long debriefing,” he said. “If you’re a principal in a building with 40 teachers, that starts to add up very quickly.” Some districts are now considering the benefits of shorter and more frequent evaluations, King said.

King concluded with a question about how he would measure his own success. “There’s this tendency to look at a project six months in, and say, Have we moved the metrics?” he replied. “But I do think that [the length of] a governor’s term is a reasonable indicator of whether progress has been made. If, in four years, we still have the same remediation rate that we do now, we will not have done our work well enough. If students aren’t staying in school and graduating ready for college and careers, then none of this matters. It’s just that simple.”

The Common Core, which is now in effect in 46 states, has also been committed to significant shifts in public school instruction. In reading, students now study a balance of informational and literary texts and are spending more time learning how to use evidence from source materials. In math, there is a new focus on coherence, fluency and application. These shifts are motivated by the skills needed for students to succeed in higher education and in the workforce. For a lot of people now, the United States has led the world in patterns, King said. But “to continue as we are, we have to hit a trail in fractions, and it turns out that fractions matter a lot more for later success in mathematics than patterns.”

Concurrent changes have emphasized improving the quality of classroom delivery. Too often, assessment has been limited to a “postmortem” analysis of a teacher’s effectiveness, King said. Data-driven instruction is a means of improving teacher effectiveness and student performance throughout the school year. As research by Jonah Rockoff, Associate Professor of finance and economics at Columbia Business School, and others have shown, a teacher’s effectiveness can have a profound impact on student achievement. Students taught by teachers who are in a series of highly effective teachers will do better, and students who have a series of ineffective teachers will do worse, King said. “And that’s why our assessments have become more challenging. The reality is that people do what is measured, and we must measure what we value.”

If meaningful performance feedback is a necessary input into helping teachers get better, evaluations must convey meaningful content, which has not traditionally been the case. Before the recent changes in the laws governing teacher evaluations, close to 99 percent of teachers were rated as at least satisfactory, according to the nonprofit The New Teacher Project. At the same time, most teachers agree that there is at least one poor performing teacher at their school. Improving the evaluation system will involve some sweeping changes, such as multiple observations by multiple observers, King said. Yet given the promising results from research on the improvements that come from evaluation-driven feedback, the benefits are worth the effort.

The state’s reform agenda is ambitious, King conceded. He is often urged to implement changes more slowly—adopt the Common Core, perhaps, but put off a new evaluation system. However, students cannot afford further delays, King argued. And the changes reinforce one another, so there is value to their simultaneous implementation. “Students are accountable for the Common Core today,” he said. “When they go to college and they’re sent to a remedial course, or they can’t get a job as a welder because they don’t understand fractions, they’re accountable for the Common Core. Every student deserves the best possible education, and our job is to deliver that.”

In the question-and-answer session that followed King’s presentation, he was asked about plans to link teacher performance to compensation. While the state has launched pilot programs to reward effective teachers, this remains a difficult area for reform. King said. “Given the history of the labor movement, with its emphasis on solidarity, it is a cultural challenge to talk about differentiating rewards,” he said. “Over time, I think that our compensation model will have to change if we want to attract the strongest possible teachers.” The state’s research has shown that it is important to link working conditions and opportunities for growth even more than compensation, he added. “Creating a performance management system in schools will lead to better working conditions,” he said. “And that will not only lead people to become teachers, but to stay teachers.”

Ray Fisman, co-director of the Social Enterprise Program, asked about efforts that were tried but found to be less than successful. King discussed the struggles some districts are facing in managing the time necessary for evaluations. “Many people think that the way to do an observation is to have an hour-long conference, and then observe for an hour, and then have an hour-long debriefing,” he said. “If you’re a principal in a building with 40 teachers, that starts to add up very quickly.” Some districts are now considering the benefits of shorter and more frequent evaluations, King said.

King concluded with a question about how he would measure his own success. “There’s this tendency to look at a project six months in, and say, Have we moved the metrics?” he replied. “But I do think that [the length of] a governor’s term is a reasonable indicator of whether progress has been made. If, in four years, we still have the same remediation rate that we do now, we will not have done our work well enough. If students aren’t staying in school and graduating ready for college and careers, then none of this matters. It’s just that simple.”

Figure 1. Presentation by John King, Commissioner of Education and President of the University of the State of New York.
More than forty years of research confirm what so many of us have experienced on a personal level: a great teacher can make a lasting difference in a student’s life. However, it is difficult to measure what makes a teacher great. Tom Kane, professor of education and economics at Harvard University, discussed his efforts to answer to two-part question: How well do various methods of teacher evaluation measure what we currently understand as effective teaching, and are these methods related to student achievement gains?

Kane and his research partners, in a project supported by the Bill & Melinda Gates Foundation, analyzed four main categories of evaluation methods. The first was classroom observations by evaluators trained to identify specific practices; this involved collecting video from 3,000 classrooms and scoring it according to multiple systems. The second was student surveys. The third was pedagogical content knowledge, or how well a teacher can diagnose students’ misunderstandings or recognize good examples for concepts such as dividing fractions. The last was student achievement gains on state tests and on supplemental tests on math and language arts concepts.

Their first and most surprising finding was that assessments of pedagogical content knowledge—at least in their current form—are unrelated to teacher effectiveness. “There’s a great interest these days in developing a sort of bar exam for teaching,” Kane said. “None of the measures of pedagogical content knowledge we’ve tested are related to student achievement, and that runs counter to what many of us would expect.”

Kane’s studies showed that a reliable classroom observation method must meet two criteria: observers need to be trained and certified, and teachers must be observed more than once. In fact, having multiple observations greatly compensates for the subjectivity of the evaluation process. “You get more than double the gain in reliability if you use a second observer,” Kane said. “Unfortunately, in a lot of schools, there’s just one person taking on this task. You could increase reliability just by having a principal and a department head observing and sharing responsibilities” (figure 2).

In his study of student surveys, Kane found that feedback from students was not only a low-cost method of evaluation, but was also the most reliable of the four methods. “This wasn’t something we predicted, but it makes a lot of sense,” he said. “Even if students are not as discerning as trained adults, this method benefits from the power of averages. You get to average over 25 students in an elementary grade or 75 to 100 students per teacher in a middle school and high school.” Surveying students also helps to compensate for variation in a teacher’s practice. In a traditional classroom evaluation, a teacher might be observed once or twice; students observe a teacher for 180 days a year.

The final part of Kane’s study focused on whether teachers who were identified as more effective actually caused their students to learn more. “Many people raise the question of whether teachers who have particularly big achievement gains might just have exceptional students,” Kane said. “The only way to resolve this issue is to randomly assign students to classrooms.” In this study, Kane asked principals to create a roster for each class. These rosters were assigned randomly to different teachers within each grade and subject. The researchers predicted student outcomes based on previous measures of their teachers’ effectiveness, including observations and surveys. The researchers then compared the predicted outcomes to the students’ actual performance.

The researchers found that teachers who were evaluated as effective had the greatest achievement gains. “Within each group of teachers there’s a lot of variation, but on average we were right,” Kane said. “This suggests you can identify effective teaching with the class, and thus a sense of their efficacy. (figure 3). The Danielson instrument is a tool used by evaluators to determine a teacher’s level of engagement with the class, and thus a sense of their efficacy.

Figure 3. The Danielson instrument is a tool used by evaluators to determine a teacher’s level of engagement with the class, and thus a sense of their efficacy.

Not surprisingly, many teachers are wary of having a camera in the classroom, he said. The Best Foot Forward Project was designed to test the benefits of allowing teachers to record at least one lesson every two weeks, and then choose five lessons to submit over the course of a year. The videos are reviewed by trained observers who can identify discreet, coachable changes. In addition, students provide several rounds of feedback throughout the school year. “For feedback to be helpful, it requires frequency; otherwise, it’s just about accountability,” Kane said. “We hope that teachers will feel empowered when they see their measures of effectiveness improve.”

Kane was asked by a conference participant about response to the video project. How do you avoid making teachers feel like they are being observed all the time?”

Kane discussed the approach adopted by teachers in Hillsborough County, Florida. The teachers agreed to record a large number of videos and allow their principals to score them, as long as the teachers could select which videos the principals could review. However, outside observers were permitted to view any of the videos.

The researchers compared the scores of the videos that were selected by the teachers for review and those that were not. Although the selected videos received higher scores on average, behaviorally identified a teacher’s effectiveness. “Once we compensated for measurement error, the correlation between the selected and unselected videos was one,” Kane said. “That means that teachers who performed particularly well on their selected videos also performed particularly well on the videos they didn’t select. By giving teachers the ability to select their videos, you’re going to get a difference on the mean score, but you’re not giving up the ability to see who’s struggling and who’s not.”

Kane concluded by discussing the urgency of using these findings to move forward with changes in the evaluation process. “We’ve known for four decades that teaching practice differs, but we’ve just getting around to taking that seriously,” he said. That is why John King and others are moving quickly in adopting changes, and also why it will take time to see the results of their efforts. “We’ve got a tremendous amount of ground to make up,” Kane said. “But I can’t think of any way we’re going to make progress without recognizing the importance of evaluating and developing teachers.”
Using Social Psychology to Lift Minority Achievement

While it has seen limited application in the classroom, the field of social psychology—the study of how our social environment influences our behavior—holds great promise for helping teachers improve their students’ effort and performance. In a presentation based on her own field research, Valerie Purdie-Vaughns, a professor of social psychology at Columbia University, discussed several specific interventions that she has shown can increase minority student effort and performance, and thus reduce the achievement gap.

Purdie-Vaughns’s research explores the psychological climate encountered by minorities and underrepresented groups. Members of these groups often face both hidden and overt biases that result in additional stresses (not faced by non-minority) that can undermine performance and motivation. This stress is known as “stereotype threat”—the anxiety of being judged according to negative biases about one’s social identity, or of behaving in a way that would inadvertently confirm such prejudices. For example, a woman from the South might feel self-conscious about her accent, because she believes it makes her seem uneducated. Or an African-American student might get nervous taking tests, out of fear that failing will reflect negatively on his race. “To increase participation among diverse groups and lower the achievement gap, we first need to alter the psychological climate and reduce this stress,” Purdie-Vaughns said.

She discussed the problem of underperformance by students who face stereotype threat: “In a typical scenario, two students, one male and one female, enroll in an Ivy League college and plan to major in math. Both have the same SAT scores, and both are capable of earning an A in their first-year math class. However, studies show that while the male student is likely to get an A, the female student is more likely to get a B. "Underrepresented groups are not getting as much out of their education,” she said. “We need to think about the social environment in which education is happening.”

For the past five years, Purdie-Vaughns has been working on Project ACHIEVE, a series of field-based interventions designed by her and other social psychologists aimed at reducing stereotype threat. Unlike previous studies that took place primarily in lab settings, this project consists of longitudinal-experimental interventions in middle schools, high schools, and colleges. All of the studies are double blind; neither teachers nor students know who is being “treated” for stereotype threat and who is in the control group.

Early research showed that there are ways to manipulate or alleviate the stresses from stereotype threat. In one study at the University of Michigan, female and male college students took a sample math GRE. In the control condition, students received standard test instructions. In the treatment condition, students were told that while the examiners were aware of gender differences in math achievement, they had uncovered no such differences tied to this particular test. “The idea is to turn off the stereotype in that moment,” Purdie-Vaughns explained. The intervention proved successful. In the control condition, male students scored three times higher than female students, in the treatment condition, the scores were almost the same. “Two things happen when we lower the heat of the stereotype,” she said. “Women’s performance is lifted significantly, and the men’s performance drops slightly.” This study and others like it have now been replicated more than 3,500 times, comparing the performance of groups such as male and female students, African-American and white students, and white male and Asian male students, with similar results.

The studies have moved out of the lab and into the classroom. Purdie-Vaughns said. One study of more than 15,000 college students found that self-reports of stereotype threat at freshman year predicted about 35 percent of the variance in grades among African-American, Latino, Native American students, and white students. Another found that self-reports by female professors of science and technology could predict how frequently they brought up their own research when speaking to male faculty. A new study by Purdie-Vaughns on cortisol, a steroid hormone linked to obesity, found that self-reports of stereotype threat by college students at the beginning of freshman year predicted a 1.0 pound weight gain by June.

As seen in the cortisol study, stereotype threat is a multi-level physiological phenomenon. “At the neurological level, stereotype threat has been associated with low recruitment of areas that are associated with learning,” Purdie-Vaughns said. “The parts of your brain that are used for learning shut down, while at the same time, the parts that deal with regulating emotions are over-engaged. Your heart beats a little faster, almost like you’re walking on a treadmill. You have a heightened immune response, which increases your chances of getting sick. And you experience a plummet in trust and in your sense of belonging, and a heightened perception that your environment is threatening.”

A common misperception is that stereotype threat is rooted in the individual, not in the environment, she noted. “Many people think that stereotype threat is something the student can meet on their own, or that the teacher set a standard for what the student should be able to achieve.”

After receiving this feedback, the students were asked if they wanted to revise their essay. In all three conditions, the amount of time that white students spent on revisions was the same. Among African-American students, those who received the positive note spent more time on revisions than those who received the unbuffered criticism, and those who received the Wise feedback spent the most time. “For minority students, these slight shifts in criticism have a strong effect on motivation,” Purdie-Vaughns said (figure 4).

Purdie-Vaughns and her research partners replicated this finding in a study of seventh-grade middle school students. The school’s population was largely middle-class, and evenly balanced between African-American and white students. The school’s population was largely middle-class, and evenly balanced between African-American and white students.
white students. All students were asked to write essays about their personal heroes, and the essays were graded by their teachers. The researchers then attached notes with either control condition or Wise feedback to the essays before they were returned to the students.

Like the college students, the seventh graders were given the opportunity to revise their work. The researchers found that the Wise feedback completely eliminated the gap in effort by African American and white students in revising their papers (figure 5). A separate study confirmed that this extra effort paid off with higher grades. “Raising classroom standards might not close achievement gaps if they don’t address stereotype-based ambiguity,” Purdie-Vaughns said. “Wisely strategies can disable students of this ambiguity and unlock motivation.”

She has also tested the effectiveness of a different method, known as value-affirmation interventions. In this method, students are asked to think about their values in a 15-minute essay about their values. The exercise, which was given about six weeks into the school year, had no effect on the end-of-quarter grades for white students. But for minority students, it significantly raised performance, reducing the achievement gap by about 35 percent.

These findings were so positive that Purdie-Vaughns and her partners replicated the study twice, with consistent results across eight schools. They also found that the benefits of value-affirmation interventions last from one year as long as three. “These interventions are not going to teach students who can’t read to read,” she said. “But sensitivity to the environment of minority groups can have a powerful effect on their performance and motivation.”

After her presentation, Purdie-Vaughns was asked if these methods work if the subject is aware of the intention. Her research has shown that knowledge of the process eliminated the effect, she answered. “No one wants to feel ‘fixed up,’ especially students,” she said. “Knowing that someone is trying to ‘fix’ you would actually raise ambiguity about your identity.” Resolving this issue is the next step in developing effective interventions. “We know that these methods work,” she concluded. “Now the question is how to scale up the results, possibly by integrating these interventions into the classroom.”

Putting Teacher Training Into Practice

In 1991, Minnesota became the first state to allow charter schools. Under the new law, educators and activists could seek permission to operate publicly funded schools outside the management of public school districts. To many reformers, this presented an opportunity to bring innovative practices and strategies to low-income communities, many of which were served by failing schools. When Trenton followed St. Paul’s move several years later, Norman Atkins, who recognized the urgent need for better schools through his work at the Robin Hood Foundation, established one of the state’s first charter schools, North Star Academy of Newark. It would become one of the highest-performing schools in the nation.

The success of North Star Academy led to the formation of Uncommon Schools, a nonprofit charter school management organization that operates some of the most successful public schools in Newark, Brooklyn, upstate New York, and Boston. But, in working to scale Uncommon Schools’ impact, Atkins was struck by how difficult it was to identify, recruit, and train teachers for the Uncommon School network. To develop an enduring pipeline of teachers, he co-founded the Teacher U at Hunter College program, now Relay Graduate School of Education (GSE), a pioneering, independent institution of higher education aimed at preparing teachers with practical skills and techniques they can implement immediately to help students succeed.

Relay GSE’s approach reflects the core belief that great teachers are not born, but made, and that a succession of highly effective teachers can change a student’s life outcome. “It takes more than one teacher to make a difference in the lives of students,” Atkins said. “It takes a ‘relay’ of great teachers.” Studies have shown that even an average student who is taught by a series of four highly effective teachers will move from the 50th percentile of students to the 90th percentile by the end of the fourth year based on standardized test scores in mathematics.

Conversely, an average student who is taught by four ineffective teachers will drop to the 37th percentile. “And that unfortunate scenario happens too often in both urban and rural schools,” Atkins said. “Those students can kiss their dreams of going to college goodbye (figure 6).”

Relay GSE’s model gives novice and early career teachers hands-on training as they are developing their skills. “We are educating people who are essentially working as full-time teachers while they are learning their profession,” Atkins said. “It’s probably not the ideal paradigm, but it’s one that works, given the economics that we face.” By teaching specific, concrete strategies, and by practicing and refining these techniques, Relay GSE hopes to better prepare teachers for the challenges and intricacies of the classroom. “The sad truth is that the vast majority of the teachers who enter the profession find themselves terribly unprepared,” Atkins said. “They struggle around the very basic moves, like how to set up their classroom and how to manage it.”

During his presentation, Atkins conveyed Relay GSE’s vision for the wide range of responsibilities teachers must have, and the role Relay GSE plays in developing these competencies. Rather than just focus on the particularities of their students, teachers need to understand themselves, including their own backgrounds and educational experiences, Atkins said. At Relay GSE, teachers learn how to engage parents and members of the community, as well as the importance of developing their students’ life skills, such as persistence and social intelligence. “At every instance, a teacher is always working on two objectives,” Atkins explained. “One might be how to convert decimals into fractions. But another is imparting zest and grit and love. Teaching character is an essential part of our program.” This doesn’t mean that
teachers should focus on pedagogical skills at the expense of subject knowledge, he added. “Under the Common Core, the bar has been raised,” he said. “It’s not sufficient anymore for teachers to just control their classrooms. They must be able to teach the content.”

The field of education faces significant, persistent obstacles in pursuit of its broader objective of elevating the status of the teaching profession, Atkins said. Studies have shown that the majority of new teachers were among the weakest students at their undergraduate schools. And about half of new teachers leave within the first five years—a trend that has pushed the average number of years of experience for U.S. teachers into the low single digits. “We need to think collectively about how we ramp people into the profession,” Atkins said. “Right now, we’re taking the good-hearted souls who want to devote their lives to teaching and we put them in the most difficult circumstances.” A growing body of research finds that first-year teachers often receive insufficient support and feedback, and they are frequently assigned to schools and classrooms with the most challenging students. In addition, many are struck by the dissonance between what they learned in their preparation programs and what their principals and districts expect them to do. At times, none of these factions seem to share goals or even a philosophical framework. “This is not the way to get people to last in the profession,” Atkins said. “Or to be happy.”

In the question-and-answer session, Atkins was asked if Relay GSE and other schools of education could better align their goals with principals and district leaders. He responded by stressing the need for meaningful evaluation and consistent feedback, so that teachers, principals, and teacher preparation can strengthen their connections and alignment.

“If we’re being honest about what’s out there, we have more than 3 million teachers who have closed the door to their classrooms, and they’re feeling incredibly lonely and vulnerable,” he said. “And their principals are dealing with the complex problems facing their schools, and are hidebound by regulations to go into classrooms only a few times during the year, and to give a standard evaluation with low-quality feedback. To change that, we need to create a generation of instructional leaders who are peripatetic in their schools, and who give teachers helpful, concrete feedback.”

Evaluations, Atkins stressed, are not a cudgel to quash spirits, but rather a vehicle to help teachers improve and succeed. He concluded by noting that the majority of teachers, when asked, say that they want to improve. They naturally want to see an upward trajectory in their own careers, and to improve and achieve, just like their students. “We have the responsibility to make this continuous improvement possible for our teachers,” Atkins said. “And if we don’t, we’re never going to be able to improve outcomes for the 50 million kids in this country.”

Atkins urged the field to create more meaningful ways to provide aspiring teachers with opportunities to develop their skill set before becoming full-time educators, whether through working as tutors or offering after-school, small-group instruction. “There are huge economic implications to these efforts,” he said. “But my hunch is that we could build a cycle in which teachers stay longer. We need to think collectively about how we on-ramp people into the profession, “ he said. “But my hunch is that we could build a cycle in which teachers stay longer. No, that’s not what we need to do it, and that we shouldn’t be afraid!”

And in these discussions, Aronson said, it is important to recognize that everyone involved in the debate on education, from reformers to principals to teachers themselves, want teachers and their students to succeed. That is what drives the reform effort and the push for measurement and feedback.

“The significance of how we train our teachers is enormous,” Aronson said. “Because when students have great teachers, we remove the limits on their abilities.”