THE HAMILTON PROJECT

Harnessing Technology to Improve K–12 Education

Aaron Chatterji
Duke University

Benjamin Jones
Northwestern University

SEPTEMBER 2012

NOTE: This discussion paper is a proposal from the authors. As emphasized in The Hamilton Project's original strategy paper, the Project was designed in part to provide a forum for leading thinkers across the nation to put forward innovative and potentially important economic policy ideas that share the Project's broad goals of promoting economic growth, broad-based participation in growth, and economic security. The authors are invited to express their own ideas in discussion papers, whether or not the Project's staff or advisory council agrees with the specific proposals. This discussion paper is offered in that spirit.

BROOKINGS

Abstract

Technological progress has consistently driven remarkable advances in the U.S. economy, yet K–12 education sees little technological change compared to other sectors, even as U.S. K–12 students increasingly lag behind students in other nations. This proposal considers how we can take a signature American strength—innovation—and apply it to K–12 education. We argue that the advent of Common Core State Standards (CCSS) and broadband Internet create promising opportunities for developing new learning technologies but that a fundamental obstacle remains: the effectiveness of learning technologies is rarely known. Not surprisingly, when no one knows what works, schools are unlikely to buy, and innovators are unlikely to create. Our proposed EDU STAR system will solve this problem by (a) undertaking rapid, rigorous, and low-cost evaluations of learning tools and (b) reporting results to the public. Coupling Internet-based real-time evaluation systems (demonstrated daily by many leading companies) with trusted reporting (modeled by *Consumer Reports* and others), the proposed EDU STAR platform will help schools make informed learning technology decisions and substantially reduce entry barriers for innovators. EDU STAR will bring together K–12 schools, teachers, and innovators and continually improve this critical foundation for economic prosperity.

Table of Contents

Chapter 1: Introduction

It is in primary and secondary schooling that individuals and nations lay key foundations for future economic prosperity. Unfortunately, recent statistics suggest that K–12 education in the United States is falling short, especially in comparison to other developed countries. We argue that educational technologies hold significant promise for improving K–12 educational outcomes, but their promise remains clouded in a marketplace where schools rarely know what works and innovators cannot easily establish the value of their products. This paper outlines a potential solution. We focus on the creation of EDU STAR, an evaluation and reporting organization, to improve the market for learning tools and create the breakthrough innovations that can continuously improve our nation's schools.

A. AMERICAN STRENGTHS AND THE K–12 CHALLENGE

The future economic growth and international competitiveness of the United States depend on successful investments to improve worker productivity and create high-quality jobs. Economists typically identify three key ingredients to economic growth: human capital, physical capital, and innovation. America's traditional strengths on these dimensions can substantially explain why Americans have long enjoyed standards of living that lead the world.

First, U.S. innovators have long created many of the world's new industries, from computers to the Internet, from automobile assembly lines to aerospace, from medical devices to biotechnology. Although other nations are making impressive gains, the American economy is still the most innovative in terms of patents, basic and applied research, and entrepreneurship. Two-thirds of Nobel Prize–winning research since 1980 has been performed in the United States, and several of the world's newest industries—such as gene sequencing, mobile apps, and social networking—see U.S.-based companies and entrepreneurs in the lead.

Second, America is traditionally the home of strong capital markets, able to allocate financing in promising directions. In addition to deep and liquid equity and debt markets for established companies, the United States excels at early-stage financing, with angel investors, venture capitalists, and others bringing capital resources that take promising, small firms—like Intel, Microsoft, Google, or Facebook—and make them global leaders in remarkably short periods.

Finally, America's university system is the envy of the world, able to attract the top minds from around the globe to its faculties and to provide students with cutting-edge higher education. These ingredients—innovation, capital, and skills—are a chain that, for a long time, has sustained American growth and advanced Americans' prosperity.

In the midst of this success, however, the U.S. K–12 education system appears to be an increasingly weak link. On quality dimensions, U.S. students' performance now lags substantially behind the rest of the industrialized world in reading, math, and science, according to tests like the Program for International Student Assessment (PISA). In terms of sheer quantity, 25 percent of American students fail to graduate after four years of high school, with many urban districts showing dropout rates near 50 percent. Given the central importance of education to economic well-being, surveys of Americans consistently list education as a top concern; education issues loom large in nearly every political campaign at the local, state, and federal levels.

Despite broad attention to education, however, the United States sees little research and development (R&D) in the K–12 education sector. Overall, 2.9 percent of total final expenditures in the United States are spent on R&D (NSF 2012). Yet in K–12 education, R&D accounts for only 0.2 percent of expenditure—one-fifteenth the average rate in the economy and one-fiftieth the rate seen in highly innovative sectors.[1] Even in a highly regulated industry such as pharmaceuticals, which happens to have total expenditures similar to U.S. K–12 education (approximately $600 billion per year), private R&D investments as a share of the total expenditure are one hundred times what we see in K–12 education.

In transportation, health, information technologies, and many other sectors, innovators and entrepreneurs radically improve services over time through investments in better products and methods. Why does the same imperative not drive innovation in education? Put another way, why does the United States fail to apply one of its signature strengths—innovation—to overcome one of its critical challenges—K–12 education?

B. BUILDING A K–12 INNOVATION ECOSYSTEM

In this paper, we first identify core challenges that foreclose innovative opportunities in educational technologies. We argue that certain features of the K–12 education marketplace make it very difficult for would-be innovators to reach their ultimate consumer—the students and teachers in our nation's schools. In particular, it is traditionally very difficult to convince a large block of these potential consumers that a given product is worth adopting. On the supply side, educational products rarely offer verifiable proof of effectiveness, creating a difficult sales proposition. If no one knows what works, how can the market work? Naturally, schools are reluctant to invest in unproven technologies. On the demand side, heterogeneous school systems and curricula traditionally limit and obfuscate demand for any given educational tool. These barriers make innovative and entrepreneurial investments unlikely to be profitable, dissuading such investments and leading to our current and traditional equilibrium where very little innovative activity is seen.

This paper proposes specific institutional innovations that will lower entry barriers, knitting together an innovative ecosystem of innovators and schools, where educational technologies will be increasingly improved and, when they are effective, adopted. Our proposal starts by noting important changes have come to K–12 education, providing important opportunities that did not exist in years past by *scaling demand*.

- First, forty-five states, the District of Columbia, and three U.S. territories have recently agreed to a set of "Common

BOX 1.

Case Studies: Into the Unknown

Since there is so little evidence for what really works, schools that do adopt educational technology not surprisingly have mixed track records. Judging by the few rigorous evaluations available, schools pick both technologies that work and those that do not. For example, a rigorous evaluation by Barrow, Markman, and Rouse (2009) found that a set of popular algebra programs called I Can Learn had significant positive impacts on math scores. On the other hand, another rigorous evaluation by Rouse and Krueger (2004) found that a popular reading program, Fast ForWord, which was being used to improve the reading skills of more than 120,000 students, did not actually have significant impact on reading skills. These evaluations are rigorous and informative, and they show that what works and what is popular can be far apart. But there are currently very few of such rigorous studies. One-off evaluations like these are slow and costly to set up and perform, with the results coming years later, at which time the technologies themselves may be outdated.

More generally, schools make technology decisions not only with little evidence to guide them, but also with few clear lessons to impart to other schools. For example, Liverpool Central Schools was one of the first school districts in New York State to experiment with educational technology, providing a laptop for every student in 2000. By 2007, however, the school announced it would discontinue the policy, citing that there was no evidence that the program actually raised educational outcomes. Meanwhile, the Mooresville Graded School District in North Carolina has experienced solid gains since launching their laptop program and associated innovations in 2009, with increased graduation rates and increases in the percentages of students who are proficient in math, reading, and science. In contrast to Liverpool, administrators, teachers, parents and students in Mooresville appear to believe that the laptops make a difference.

Two laptop programs, two school districts, two different results. This is not surprising. In neither case is there an evaluation mechanism that could separate the effects of technology investments from the many other forces at work in these school systems, or identify what features of these programs work well or badly. The lessons one can draw, or apply in other schools, are unclear.

With conflicting but ambiguous experiences, few rigorous trials, and no way to solve this uncertainty in a low-cost or timely manner, investing in educational technology is risky for both schools and innovators, and does not put K–12 education on any clear path to greater success. We are unlikely to unlock the true potential of educational technology following the path we are on. EDU STAR aims to solve this fundamental proof of effectiveness problem.

Core State Standards (CCSS), creating large-scale demand for skill acquisition on well-defined, specific tasks. These standards aggregate and clarify demand, giving would-be innovators and entrepreneurs clear, large targets for education technology tools.

- Second, the diffusion of broadband Internet platforms through schools creates another type of standard—the student-technology interface—that further scales and unifies the marketplace on a platform with high-performance capabilities while also lowering development and distribution costs for learning tools.

Both of these developments help improve innovation incentives. However, neither speaks to a more basic challenge, which remains a critical obstacle to the effective creation and diffusion of superior educational technologies: *knowing what works*. It is often not obvious whether a given educational technology improves outcomes, and formal assessments often refute popular claims. This uncertainty leaves buyers both (a) less willing to make purchases; and (b) when making purchases, more likely to choose technologies that are not effective and potentially worse than the status quo. Box 1 provides examples to highlight this basic challenge.

The core of this paper details a further institutional innovation—the missing piece—that we believe is essential to building a vibrant and effective K–12 education innovation ecosystem. We aim to solve the *proof of effectiveness* challenge through a simple, new platform that connects schools and innovators to provide rigorous, low-cost, trusted evaluation and reporting on educational technology innovations.

C. SUMMARY OF PROPOSAL

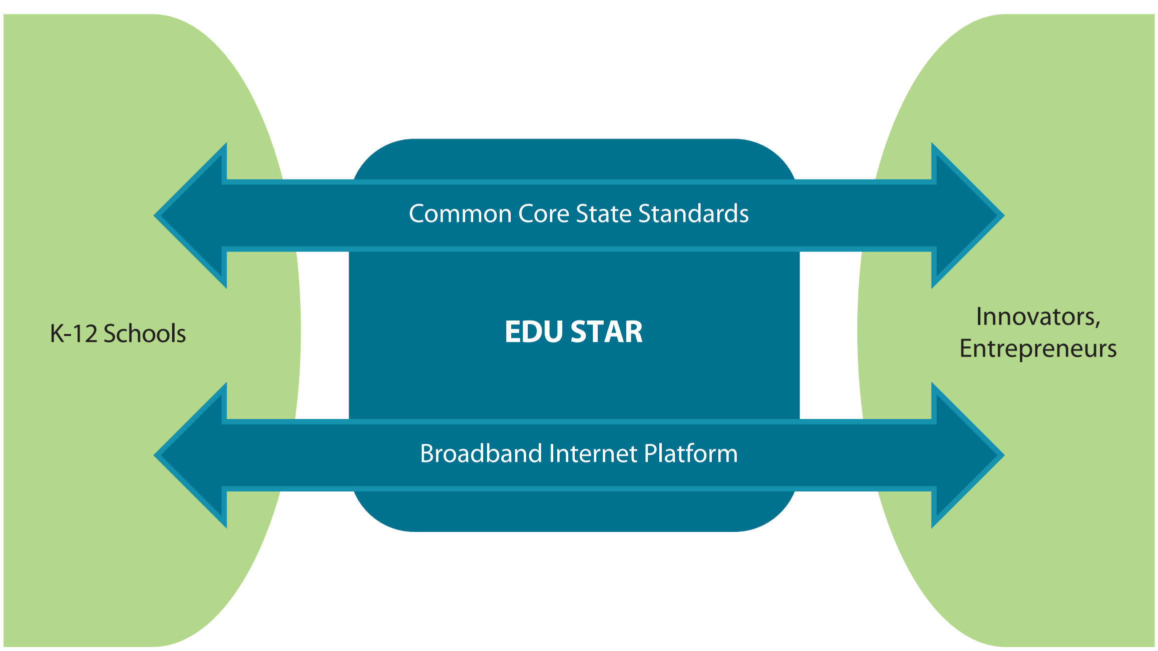
The specific recommendations outlined in this paper focus on digitally delivered learning tools in the K–12 setting, where we believe there is the greatest need and the lowest-cost opportunities, but the proposed system can naturally be extended to other types of educational innovations.

We propose the creation of a new organization, EDU STAR, to (a) evaluate and (b) report on learning technologies, thus building a trusted, rigorous, and low-cost platform that connects innovators to schools.

As a 501(c)3 nonprofit organization, EDU STAR will act as the connective tissue between innovators and entrepreneurs (suppliers) on one side and school systems (demanders) on the other. It will create a commonly shared, standardized platform that will provide two primary services.

- First, EDU STAR will perform low-cost, rapid, and rigorous testing of educational technologies by implementing real-time randomized-controlled trials of learning tools in schools.

FIGURE 1.
The EDU STAR System



- Second, EDU STAR will communicate technology evaluation results, both to the school community at large, so they can assess what works, and to the innovators and entrepreneurs themselves, so they can further improve their products.

Figure 1 provides a schematic of how EDU STAR sits on top of curricular standards (the CCSS) and a technology platform (the Internet). As innovators develop new digital learning technologies, EDU STAR will test their success against specific CCSS skills, reaching schools over the Internet platform. For the randomized-controlled trials, think Google (rapid, nearly zero-cost trials) rather than the Food and Drug Administration (years-long, extraordinarily costly trials). In particular, EDU STAR will do what many of the world's leading companies do every day: randomized testing of their business methods across their user base. Here the user base is provided by a network of participating schools, where students test software during scheduled computer time.

By coupling rigorous, transparent evaluation with active dissemination of its findings, EDU STAR also serves as a trusted reporting organization for the education technology marketplace. Schools can now buy learning tools based on definitive evidence. Innovators learn quickly whether their software or other digital learning tools are effective, and then use that information to make continuous improvements.

In short, we propose a simple but essential institutional layer for creating a vibrant, innovative, and effective education technology marketplace. We argue that this system is essential to overcome traditional obstacles that impede innovation in this sector of the economy.

Let us be clear that this proposal does not aim to encompass all forms of education technologies. The opportunity of the CCSS is to unify demand around learning objectives, and the opportunity of broadband Internet is especially conducive to digitally delivered content. Thus, our proposal primarily emphasizes a targeted but important area: digitally delivered learning tools. Moreover, the system will initially emphasize short-run interventions, testing whether a given learning tool can engage students and improve their success at a CCSS skill in real time. More generally, the evaluative architecture we propose will also allow substantially lower-cost evaluations of other types of technologies and interventions, including longer-term interventions. This proposal also considers these further opportunities.

The rest of this paper is organized as follows. Chapter 2 reviews K–12 education in the United States and the challenges it has posed for educational technology innovation. We treat the design and implementation of our EDU STAR proposal in detail in Chapter 3, including discussion of its essential parts, potential sponsors and partners, priorities, and timelines. Chapter 4 considers potential challenges to the effective implementation of the EDU STAR system. Chapter 5 concludes.

Chapter 2: The Status Quo: Challenges and Opportunities

This section first briefly reviews evidence on the challenges facing K–12 education in the United States. We then consider at greater length the lack of innovation in the K–12 sector relative to other industries. Using a simple framework for understanding healthy innovation ecosystems, we diagnose reasons for the relatively weak innovative and entrepreneurial investments seen in educational technology, ultimately obstructing the development of innovative educational tools and decoupling the K–12 education sector from the powerful creative engine that drives progress in so many other sectors. These diagnoses motivate our specific proposal, introduced in Chapter 3.

A. K–12 EDUCATIONAL OUTCOMES IN THE UNITED STATES

K–12 education in the United States lags behind most industrialized nations according to a variety of different statistics. The most noted comparison between the United States and OECD countries comes from the Program for International Student Assessment (PISA), which tests fifteen- and sixteen-year-olds from participating countries on math, science, and reading. In table 1, we present the results from the most recent PISA, in 2009. Among the thirty-four developed economies, the United States ranks seventeenth in reading, twenty-third in sciences, and thirty-first in mathematics—while Singapore, Korea, and Hong Kong rank in the top five, on average (OECD 2010).

This poor performance is also reflected in a high dropout rate. On average, one-quarter of all freshmen in American public high schools have not graduated four years later—a rate that rises toward four of every ten students in some individual states and above five in ten for some of the nation's largest school districts (EPE Research Center 2010; Stillwell, Savle, and Plotts, 2011). These statistics are even more troubling when we account for the fact that the United States spends more money per student than almost any country in the industrialized world (OECD 2011).

Many experts are concerned that the challenges in U.S. education have worsened over time. Keeping in mind that K–12 education is the foundation for further education, poor performance of primary and secondary educational services not only has wide effects for the secondary-educated U.S. workforce, but also has potential spillover effects for higher-education preparation and attainment. For example, according to work by Claudia Goldin and Lawrence Katz, the average years of schooling among Americans rose consistently for approximately ninety years, before leveling off for the cohort born in the 1970s (see figure 2). Relatedly, the rate of college attainment among Americans aged twenty-five to thirty-four has slipped to tenth among OECD countries, from first place a generation ago (OECD 2009).

On average, a quarter of all freshmen in American public high schools have not graduated four years later—a rate that rises toward four of every ten students in some individual states and above five in ten for some of the nation's largest school districts.

B. THE POTENTIAL OF INNOVATION

These facts and international comparisons suggest that the U.S. K–12 educational system has ample room to improve. While there are numerous recommendations for arresting the decline in American K–12 education, we focus on how educational technology can improve outcomes—continuously. Many sectors of the American economy see such continual innovation, leading to remarkable improvements over time. In health, consider the advent of vaccines, antibiotics, diagnostic tools, surgical interventions, and the extensions in life expectancy they bring. In transportation, think of

TABLE 1.

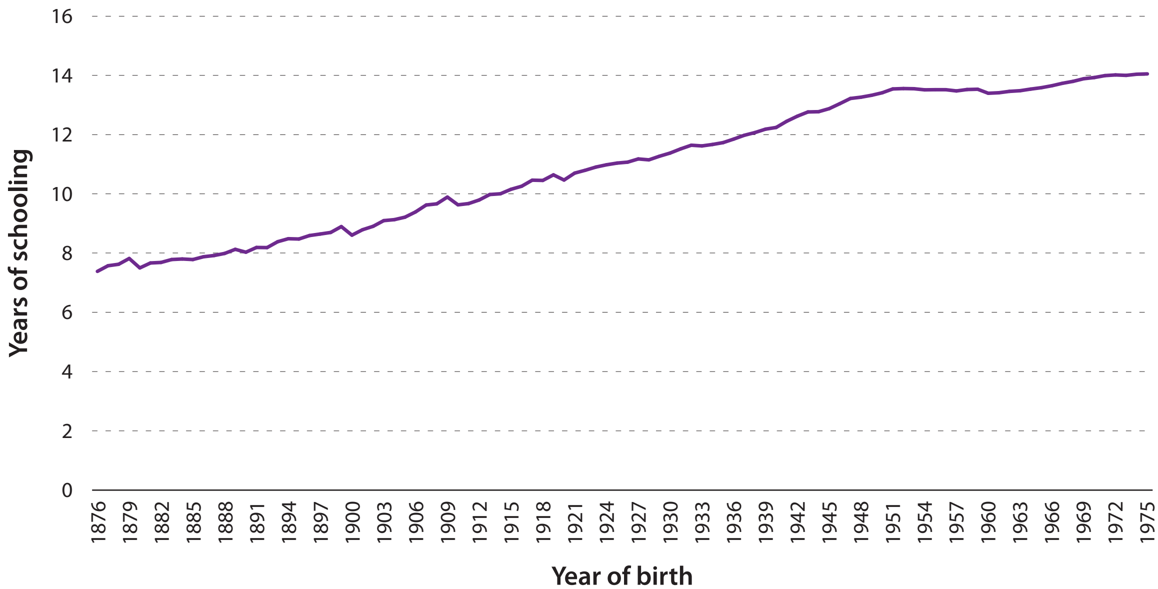
PISA Scores, 2009

Reading		Science		Math	
Location/Rank	**Score**	**Location/Rank**	**Score**	**Location/Rank**	**Score**
1. Shanghai	556	1. Shanghai	575	1. Shanghai	600
2. Korea	539	2. Finland	554	2. Singapore	562
3. Finland	536	3. Hong Kong	549	3. Hong Kong	555
4. Hong Kong	533	4. Singapore	542	4. Korea	546
5. Singapore	526	5. Japan	539	5. Taipei	543
6. Canada	524	6. Korea	538	6. Finland	541
7. New Zealand	521	7. New Zealand	532	7. Liechtenstein	536
8. Japan	520	8. Canada	529	8. Switzerland	534
9. Australia	515	9. Estonia	528	9. Japan	529
10. Netherlands	508	10. Australia	527	10. Canada	527
11. Belgium	506	11. Netherlands	522	11. Netherlands	526
12. Norway	503	12. Chinese Taipei	520	12. Macao	525
13. Estonia	501	13. Germany	520	13. New Zealand	519
14. Switzerland	501	14. Liechtenstein	520	14. Belgium	515
15. Poland	500	15. Switzerland	517	15. Australia	514
16. Iceland	500	16. United Kingdom	514	16. Germany	513
17. United States	**500**	17. Slovenia	512	17. Estonia	512
		18. Macao-China	511	18. Iceland	507
		19. Poland	508	19. Denmark	503
		20. Ireland	508	20. Slovenia	501
		21. Belgium	507	21. Norway	498
		22. Hungary	503	22. France	497
		23. United States	**502**	23. Slovakia	497
				24. Austria	496
				25. Poland	495
				26. Sweden	494
				27. Czech Republic	493
				28. United Kingdom	492
				29. Hungary	490
				30. Luxembourg	489
				31. United States	**487**

Source: OECD (2010).

FIGURE 2.

Years of Schooling in the United States by Year of Birth, 1876–1975



Source: Goldin and Katz (2007).

the transition, starting with the Industrial Revolution, from horse-drawn carriages to trains, cars, and airplanes that travel distances in hours that once took months. In major industries—from agriculture to manufacturing, from telecommunications to information technology—the story is one of remarkable and ongoing advance.

Such progress comes, ultimately, from many specific innovations and the decisions to invest in them. As one metric of these efforts, table 2 shows rates of explicit R&D as a fraction of total revenues for several innovative industries in the United States. The pharmaceutical sector, where private R&D rates hit 22.9 percent of revenues, leads the pack, but there are many notable sectors where private firms choose to devote enormous resources to such innovative activity. By contrast, estimates suggest the K–12 education sector sees no more than 0.2 percent of expenditure on R&D, suggesting that there is a problem impeding innovation investment—but also an opportunity.

The mobile apps market is a current and aspirational example of a remarkably healthy innovative ecosystem. Within two years of their introductions, the Apple and Android software platforms running on smartphones and tablets together saw more than 300,000 different software applications produced, with more than 80,000 different companies worldwide generating these applications. This enormous scale of innovation and entrepreneurship did not come by some magical force, but because large numbers of people chose to make explicit investments, attempting to produce new and better tools.

A simple and general way to understand the “innovative choice” is to consider the potential payoff to the would-be innovator. On the one hand, the innovator must consider the cost, C, of developing and marketing an innovation. On the other hand, the innovator hopes to attain a benefit, an expected market value, V, from a successful new product or service. Naturally, in industries where the expected value (V) tends to exceed the innovation cost (C), we will see lots of innovation.

For example, pharmaceuticals are an industry where the costs of developing a new drug, C, are very high, and yet the value, V, for an effective drug is so large that we still see enormous amounts of innovation investments, as noted above, and many profitable companies. For mobile apps developers, the costs of developing a new app is relatively low, whereas the market scale—knit together across common platforms (Apple’s iOS and Google’s Android)—makes the market potential, V, extremely large. Hence, we see 80,000 different companies innovating in this space only two years after the platforms were created.

In contrast, educational technology appears to lag far behind. While K–12 educational sector expenditure in the United States—$625 billion in 2007–2008—is similar in scale to the global market for U.S.-based pharmaceutical companies, the rate of R&D expenditure appears one hundred times smaller. The worldwide mobile apps market, with $8.5 billion in revenue in 2011, is a tiny fraction of what is spent on K–12 education. Why do we not see 80,000 different companies focused on the K–12 education space, drawing together the expertise and creativity of teachers, software designers, entrepreneurs, and investors to produce innovative learning tools? Why is education technology different?

TABLE 2.

R&D Expenditure Rates by Industry

Industry	NAICS Code	R&D/Output (%)
Aerospace products and parts	3364	8.2
Computer and electronic products	334	12.3
Machinery	333	3.1
Medical equipment and supplies	3391	5.1
Pharmaceuticals and medicines	3254	22.9
Semiconductor and other elec. components	3344	15.7

Source: BEA (2012); NSF (2011).
Note: Year is 2006. R&D expenditure excludes federal funding.

C. THE CURRENT K–12 EDUCATION TECHNOLOGY MARKET

We next consider the nature of the current K–12 education technology market, and provide a diagnosis of critical challenges that impede innovation and entrepreneurship in this area. We begin with some details on current educational technology markets before turning to an analysis of the institutional features that, traditionally, have made the expected profit look unattractive for would-be innovators. This analysis leads directly to our proposal in Chapter 3.

i. Types of Education Technology and Our Focus on Instructional Content

Education technology encompasses (a) baseline hardware, including computers, network equipment, tablets, and smart boards; and (b) content layers, including instructional software, digitally delivered textbooks and lectures, test preparation and assessment software, and enterprise software to help manage school systems, schools, and classrooms.

While all of these applications and markets are important, this report will focus primarily on instructional content. We focus on instructional content for three primary reasons, pointing to scalable, low-cost opportunities:

- First, the advent and adoption of CCSS resolves disagreements over K–12 learning objectives and provides specific objectives by subject and grade level, providing larger-scale markets and clear guidance for instructional software developers.

- Second, the spread of broadband Internet now provides a standardized, large-scale, high-quality distribution platform that can reduce the costs of market access. The Internet platform also provides new opportunities for rapid, rigorous, and low-cost evaluation.[2]

- Third, instructional software can personalize learning, adapting to the needs of individual students, and evolving as the student progresses. The capacity for personalizing learning is an important opportunity in this space, much as "personalized medicine" is an important opportunity in health care.

To date, instructional content investments have lagged, leaving a technology content gap despite substantial investments in underlying computer hardware and Internet connectivity. The hardware and software are, naturally, complements, and it is not surprising that in the absence of a strong content layer, school systems spend rather little on instructional technology.[3] According to one estimate, the entire U.S. market for K–12 instructional materials, including the textbook market, is $19 billion annually (NeXtup 2010). While this is a significant market, it is dwarfed by the total annual spending of more than $600 billion on K–12 education. Even within this small slice, educational technology spending is meager, especially in K–12, further indicating the presence of particular barriers to technology development and adoption (see the Council of Economic Advisers 2011).

ii. Market Demand Features

The U.S. K–12 school market is fragmented, consisting of many different buyers with distinct preferences and requirements. Schools are governed largely at the local level, which not only makes buyers independent in their decisions, but also typically results in divergent curricular objectives and procurement procedures. In 2007–2008, there were 98,916 public K–12 schools arranged into 13,754 different school districts, which are typically the relevant buyers for educational technology decisions. In addition, there were 33,740 private K–12 schools, acting as additional, independent buyers. Some large school districts create consolidated demand centers. For example, the largest five districts in the country collectively serve 2.7 million students and spend $35.8 billion annually.[4]

In addition to fragmentation across buyers, instructional content is typically age or grade specific, meaning that the market size for any particular learning application may only be a fraction of the overall market in a given year. Nonetheless, with more than $600 billion spent annually on K–12 education in the United States, average spending per grade level is approximately $50 billion, which is still very large. As with pharmaceuticals, which divide the market by numerous therapeutic categories, the scale of these submarkets appears more than sufficient in principle to attract substantial R&D investment.

Finally, there are after-school (tutorial) programs and household consumers, who can buy educational technologies, separate from school systems. With application downloads across iOS and Android platforms growing at more than 300 percent per year, there appears to be strong momentum and additional potential for educational technologies in this space.

While the potential market is large and there are numerous interested businesses and nonprofits, educational technology remains an extremely small share of K–12 school expenditures.

iii. Market Supply Features

Instructional content is currently dominated by the "Big Three" textbook providers. These companies—Pearson, McGraw Hill, and Houghton-Mifflin—capture 85 percent of the textbook market, selling under various brand names (often acquired companies) through consolidated marketing operations. There is a much larger set of smaller players, many with subject-specific or region-specific niches.[5] The Big Three are also active in education technology, leveraging their marketing platforms to sell such products. The Big Three are translating content into eBook format, building courseware on new platforms like the iPad, and developing software applications. The market's consolidation around a few core suppliers is consistent, as economics suggests, with large barriers to entry.

Recall that instructional software is a small market within educational technology. While precise data is difficult to obtain, our discussions with industry participants indicate that the largest K–12 educational technology segment is computer hardware (provided by well-established companies such as Dell and Apple), followed by enterprise software systems. Instructional software is estimated to be one-tenth the size of the computer hardware market. Within instructional software, one venture capitalist we interviewed estimated that there are 113 software products related to K–12 instruction, while the Software & Information Industry Association Education Division counts more than 150 companies in its membership.[6]

Early-stage financing on the supply side is correspondingly low. Estimates suggest that venture capital totaled $200 million annually for education companies in the past five years, backing an average of twenty-five new businesses per year. This venture capital investment compares to $3.7 billion for biotechnology, $2.3 billion for medical devices, and $4.0 billion for software.[7] Some observers suggest there is far more money in this space, or on the sidelines, but little willingness to buy technologies because of the downstream obstacles, which we consider below.[8]

Financing challenges also appear at exit. Industry participants observe that the typical path for successful entrepreneurial firms ends in acquisition by established companies (especially one of the Big Three) rather than scaling independently. Relatedly, IPOs for education companies are rare, averaging only 1.5 per year for the past two decades, which is one-twentieth the rate in the life sciences. The prevalence of exit by acquisition is consistent with the downstream barriers—the high costs of reaching local school systems. The advantages of incumbent downstream players also suggest their strong bargaining positions, which may reduce the ultimate payments to entrepreneurs and dissuade entry.[9]

D. CORE OBSTACLES TO EDUCATION TECHNOLOGY

While the potential market is large and there are numerous interested businesses and nonprofits, educational technology remains an extremely small share of K–12 school expenditures. The persistently small size of the educational technology market suggests that its development faces serious challenges.

In particular, entrepreneurs with cutting-edge educational technologies face two immediate problems in selling to schools: (a) school systems are not easily convinced that a given technology is valuable, making it hard to sell; and (b) while the overall market is large, each school system is relatively small, with idiosyncratic procurement methods and governance. The market is thus large in total but hard to penetrate. In practice, these factors likely reduce the incentives of entrepreneurs and innovators to enter the market in the first place.

To clarify the challenge for innovators in education technology, consider again the general framework above for understanding the "innovative choice," where the innovator weighs the expected value (V) of an innovation against the cost (C) of attaining it (see chapter 2.B). For educational tools, the market value—while large when aggregated across school districts—becomes in practice a district-by-district proposition, cutting V into small pieces and raising C as the innovator engages each school system independently. Unlike pharmaceuticals that—once rigorously proven according to established standards—are taken up by myriad insurers and hospitals, there is currently no proof of effectiveness that educational technology buyers can rely on. The innovator thus struggles—and often fails—to convince any particular school district that the innovation is worthwhile. Entry into each school system becomes a separate, slow, expensive, and uncertain process, raising costs (C) above expected sales (V), and making innovative investments unattractive in the current environment.

We can further articulate the specific issues as follows:

i. The Effectiveness Challenge

Ineffective evaluation, and hence the lack of credible information about what works, is a serious roadblock to a robust market for educational technology.

- **Buyer uncertainty.** To the buyer, it is often not obvious that a given technology improves an important outcome, and educational technologies are fraught with claims and counterclaims about their usefulness. Verifying claims about a technology's benefits by running an evaluation is typically an unrealistic step for a school system to take. Furthermore, the measurability of outcomes can be difficult, effectiveness may vary by context, and educational goals themselves are not always agreed upon.
- **Seller costs.** Buyer uncertainty translates into very high marketing costs. Sellers find it costly to signal quality in a convincing manner. Undertaking traditional, one-off evaluations of their own products can be very costly and slow. Even if a seller pays for rigorous third-party evaluation, the buyer may struggle to assess the study and may not trust third parties that the seller pays.

TABLE 3.

Common Core State Standards: Examples from First Grade Mathematics

Area	Standard	Description
Operations & Algebraic Thinking	1.OA.1	Use addition and subtraction within 20 to solve word problems involving situations of adding to, taking from, putting together, taking apart, and comparing, with unknowns in all positions, e.g., by using objects, drawings, and equations with a symbol for the unknown number to represent the problem.
	1.OA.7	Understand the meaning of the equal sign, and determine if equations involving addition and subtraction are true or false. For example, which of the following equations are true and which are false? 6 = 6, 7 = 8 - 1, 5 + 2 = 2 + 5, 4 + 1 = 5 + 2.
Number & Operations in Base Ten	1.NBT.1	Count to 120, starting at any number less than 120. In this range, read and write numerals and represent a number of objects with a written numeral.
	1.NBT.5	Given a two-digit number, mentally find 10 more or 10 less than the number, without having to count; explain the reasoning used.
Measurement & Data	1.MD.1	Order three objects by length; compare the lengths of two objects indirectly by using a third object.
	1.MD.3	Tell and write time in hours and half-hours using analog and digital clocks.
Geometry	1.G.1	Distinguish between defining attributes (e.g., triangles are closed and three-sided) versus non-defining attributes (e.g., color, orientation, overall size); build and draw shapes to possess defining attributes.
	1.G.3	Partition circles and rectangles into two and four equal shares, describe the shares using the words halves, fourths, and quarters, and use the phrases half of, fourth of, and quarter of. Describe the whole as two of, or four of the shares. Understand for these examples that decomposing into more equal shares creates smaller shares.

Source: Common Core State Standards (www.corestandards.org).

Without effective evaluation, the educational sector cannot judge the best products. At the same time, we cannot expect that the best products will be adopted, dissuading innovation and fundamentally limiting the ability of educational technology to improve outcomes.

Note that even for researchers, who specialize in studying educational interventions, the set-up costs and time costs for a typical evaluation are large. Researchers often must develop one-off relationships with numerous decision-makers within a particular school system, develop custom-made assessment approaches, organize and run the experiment, collect data, and undertake customized analyses. These challenges—leading to slow and costly intervention assessments—parallel the problems with innovation in this sector itself.

By contrast, the ideal evaluation and reporting system will be (a) rapid and low-cost, (b) rigorous, and (c) widely trusted. As discussed below, these design features are relatively straightforward to achieve for digitally delivered learning tools, and mirror existing evaluative platforms of many businesses.[10]

With a convincing, market-wide stamp of approval, the innovator substantially lowers the marketing costs in each district while increasing the probability that a given district perceives the product as valuable. Currently, in the absence of clear evaluation of technologies, the barriers to entry for innovators rise, raising costs for all innovators and entrepreneurs.

ii. The Procurement Systems Challenge

Given the lack of evaluation and reporting, local school systems have little idea what to buy. Not surprisingly, procurement systems in school districts become complex, idiosyncratic, and slow, and effectiveness itself, which is so rarely known, does not in practice drive choices. Technology providers must in turn go to great lengths to convince a given set of local decision-makers to devote limited budgets to any particular technology. Observers note several particular aspects to the procurement challenge that raise large entry barriers and are consistent with the innate ambiguity about what works, including the following:

- **Relationship-based sales.** Sales within a single school district can require substantial contact with numerous influential people.[11] Given these many points of contact, which differ from district to district, salespeople cultivate long-standing relationships with local decision-makers. These ties reduce the cost of marketing each new round of products (for the company that pays the specific salesperson), but make sales expensive and tilt decisions away from entrepreneurs.

- **Lack of buyer training.** District administrators may have limited training in business, procurement, or evaluation, and some observers argue that school decision-makers typically pay little attention to the return on investment. Furthermore, only 51 percent of schools have a full-time technology coordinator.[12]

- **Inflexibility of discretionary funding.** School systems not only allot very little expenditure to educational technology, especially instruction, but also do not pivot their funds quickly. Berger and Stevenson (2007) report that the sales cycle takes eight months to secure schools in a district and eighteen months to expand across a district.

These challenges help explain the market concentration in education technology, where large, patient suppliers with existing contacts and sales networks are able to introduce products more broadly and quickly, while would-be entrants face especially large entry barriers. Moreover, because these sales contacts are also time-consuming for the local buyer, local schools may favor purchasing a suite of products from a single company rather than multiple products from different companies and smaller entrants.

By providing trusted and rigorous evaluation, our proposal can substantially limit the district-by-district marketing expenses that technology sales otherwise require. In essence, the platform we propose will substantially reduce entry barriers.

E. KEY POINTS OF PROGRESS

Recent progress in developing CCSS helps resolve disagreements over learning objectives, focusing on mathematics and English language skills by grade level. For example, the CCSS first grade mathematics standards include twenty-one specific skills (see table 4 for excerpts). These standards are highly specific, providing clear, actionable guidance for entrepreneurs and businesses in developing learning technologies. With forty-five states, the District of Columbia, and three U.S. territories adopting these standards, there is now substantial scale in the marketplace for each of these specific skill objectives.

Furthermore, for digitally-based learning technologies, the Internet now provides a large-scale, low-cost, high-function distribution platform that can reduce the costs of market access and allow for low-cost evaluation. It further lowers distribution and development costs for software or other instructional content innovators. It also can help schools avoid customization problems where they end up being locked in to the products of a single firm. The spread and advance of ICT (information and communications technology) infrastructure thus provides an important opportunity, bringing educational advances into reach.

Given these developments, the time is ripe for substantial advances in educational technology solutions, but significant challenges remain.

Chapter 3: The Proposal

A. WHAT IS EDU STAR?

We propose a simple institutional innovation, EDU STAR, that can unleash the potential of instructional technologies, creating a marketplace where the best technologies are recognized and continuously improved.

EDU STAR will build a bridge between innovators and entrepreneurs, and K–12 schools. Using this bridge, EDU STAR will both (a) evaluate and (b) report on education technologies, creating a trusted, rigorous, and low-cost platform to

- Resolve uncertainties on the buyers' side of the market,
- Focus sales on the best technologies,
- Substantially reduce entry costs for new ideas, and
- Expand and accelerate the development of innovative technologies by established firms and entrepreneurs.

To be clear, our proposal is neither the Apollo space program nor the Food and Drug Administration.[13] Instead, we propose an information technology "consumer reports" for learning tools, a light-touch institutional innovation that acts as a platform for a vibrant innovation ecosystem. As a metaphor, we want to make entry costs low enough for K–12 instructional tools so that the explosion of innovation and entrepreneurship seen from the advent of the Internet can be directed into this critical space for the U.S. (and global) economy.

To be successful, the evaluation and reporting system should be

- **Rigorous.** Evaluation methods (e.g., rapid randomized controlled trials) must provide convincing evidence of success.
- **Low-cost.** Building ready test beds over low-cost distribution platforms will reduce evaluation costs towards zero, facilitating entry.
- **Widely trusted.** A third-party evaluator that schools know and trust, reporting on effective technologies, allows buyers to choose technologies without costly independent verification. It is important that this trust rely on rigorous evaluation.

To assist effective evaluation, the technologies should emphasize noncontroversial and easily measurable outcomes and CCSS (e.g., vocabulary, grammar, specific mathematical skills). The CCSS take just such a focus, which is another reason that EDU STAR will build on them.

Numerous organizations across diverse markets provide ratings and reporting. Examples include *Consumer Reports* for consumer products, and the U.S. Green Building Council's Leadership in Energy & Environmental Design (LEED) rating system for assessing whether buildings are environmentally sustainable. Consider that *Consumer Reports*, as a trusted third party that provides easily digestible assessments directly to consumers, not only helps consumers improve their choices, but also helps encourage entry in the product space by allowing small companies with good products to better establish themselves.

Meanwhile, rigorous evaluation plays a critical role in supporting many product markets, whether through the business strategies of individual corporations (see, e.g., Manzi 2012) or through regulatory bodies. For example, compare the market for pharmaceuticals, where effectiveness is proven through rigorous FDA trials. When successful, insurance companies and hospitals (think: school systems) and doctors (think: teachers) tend to adopt proven drugs at scale—even though doctors have many ideological debates and health-care systems are notoriously balkanized. Rigorous evaluation to prove efficacy can thus be critical to facilitating adoption and incentivizing R&D. This example suggests that even balkanized markets will adopt proven best practices.

It is important to emphasize that for digitally delivered learning tools, student evaluation can be embedded automatically, allowing extremely low-cost and real-time technology assessment, following the same assessment models as many large corporations.

B. AN EDU STAR EXAMPLE

To make this proposal more tangible, imagine a classroom full of first graders in Binghamton, New York, trying to learn to solve word problems involving addition and subtraction up to the number twenty. This competency is part of the CCSS (specifically, it is standard 1.OA.1) that most of the United States has now agreed to. There are many different ways to approach such problems and a teacher may want to use educational software to supplement the lesson. How would the teacher know what software to choose? Would the teacher have any statistics on whether any software program does what it is supposed to do?

Far away, in Boulder, Colorado, there is a little software company with top-notch software designers who have successfully developed some engaging online games. They meet with a celebrated local teacher who has created great methods for teaching elementary-age children how to solve math problems. The teacher thinks that key practice exercises he developed over his career could be meaningfully implemented as software tools. The software company staff is intrigued but, naturally, very nervous about investing limited time and money in such a product. They do not know about that classroom in Binghamton, or others; they do not have relationships with school administrators who make purchasing decisions; and even if they worked hard to convince a few schools to try it out, they are doubtful they could recoup the marketing costs, let alone the costs of the software development.

EDU STAR solves this disconnect between potentially great ideas and students who could benefit from them. EDU STAR would evaluate software against particular outcomes (i.e., How well does this software help children improve their skills at addition and subtraction up to the number twenty?) and report these results to the public. The leading software applications across different subjects and skills, whether they came from big established firms like Pearson or an entrepreneur in Boulder, Colorado, would receive trusted assessments and valuable recognition, allowing teachers all around the country to find and use them.

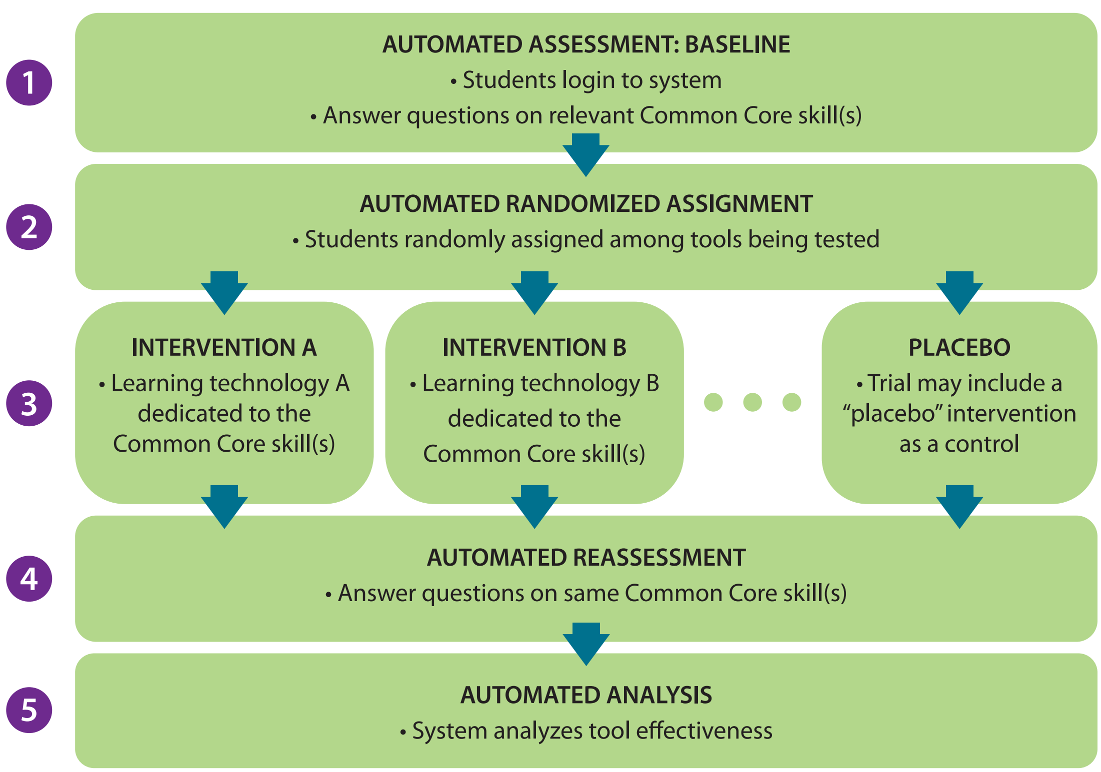
C. THE CORE FUNCTIONS OF EDU STAR

EDU STAR will perform four core functions.

- First, EDU STAR will link with market suppliers. It will screen educational software to ensure that software tested in schools meets basic design criteria. The organization will also iterate with innovators as they reengineer products based on evaluation results.
- Second, EDU STAR will link with schools, forming a permanent test bed for evaluation. This "school-facing" component for EDU STAR involves commitments from schools to use portions of their students' computer time to help evaluate promising software.
- Third, EDU STAR will evaluate instructional content by running automated randomized-controlled trials on the test bed. Building on best practices of companies like Amazon and Google, the evaluation piece will include the maintenance of computer servers and implementation of automated algorithms that integrate pre- and postassessment with randomized assigning of treatment and control groups during student computer sessions, and

FIGURE 3.

The EDU STAR Evaluation System



automated (real-time) evaluation of these randomized controlled trials.

- Finally, EDU STAR will report on education technologies by disseminating the results of evaluations and indicating the highest-performance digital learning tools for each specific Common Core standard, publishing these results in easily digestible formats online. These reports would include insights about heterogeneous treatment effects—information on whether the software tool is more or less effective depending on the characteristics of the student population, such as their starting point with the particular skill.

D. THE IMPLEMENTATION OF EDU STAR

We next consider the implementation of these core functions, beginning with evaluation.

i. The Evaluation System

Figure 3 provides a schematic of the EDU STAR evaluation system. The first step in the evaluation system will be for students to log in and complete a short assessment to establish a baseline estimate of their skills in the relevant Common Core areas. The student login would establish a unique ID to track progress within and across each Common Core area. All unique IDs will preserve student anonymity. Many schools already set aside specific class periods for computer work, and some schools, like the School of One in New York City, integrate computer-based assessment into every school day.

As an example, take a third-grade CCSS that requires students to "know and apply grade-level phonics and word analysis skills in decoding words" (Standard RF.3.3). At the beginning of class, a student might be asked to identify the meaning of various multisyllable words, irregularly spelled words, or various word prefixes (explicit subskills for RF.3.3). Fortunately, these assessments are already being developed as part of the CCSS, so EDU STAR could simply leverage existing work.

After the initial assessment, the students would be automatically randomly assigned to a software program explicitly designed to address this CCSS. For simplicity, assume there are two software applications yet to be evaluated in EDU STAR that were designed to help students master RF.3.3. Ideally, even within a class of twenty-five students, the students would be randomly assigned to one of these two software applications or a placebo intervention (involving unrelated activities).

After completing the instructional content, follow-up assessments would occur at the end of class for an immediate analysis. EDU STAR would capture the baseline performance, the instructional content used, and performance on the follow-up assessment. Aggregating the results with all of the other students working on the platform, EDU STAR could rapidly generate an evaluation of the two interventions against the placebo. Assuming EDU STAR had a large enough sample size (any single major school district in the United States could satisfy this condition), the results could literally be generated before the beginning of the next class period.

It is important to note that EDU STAR, as an automated evaluation platform, could also facilitate a broad range of interventions with far lower cost assessment than is possible today. Our primary focus is on the "lowest hanging fruit"—digitally delivered learning tools evaluated over short periods. For example,

- Software-based instructional content evaluated over a single class period, allowing immediate analysis of its effectiveness; and
- Other digitally delivered content, such as Khan Academy–style lectures, evaluated in real time.

Once the EDU STAR system is established, the platform could also be extended to evaluate longer-term interventions and interventions that are not computer based. For example,

- Many promising educational technologies are centered on helping teachers manage their classrooms or share lesson plans. These technologies, like EdModo and ClassDojo, could also be evaluated rigorously and at much lower cost using the EDU STAR platform.
- Skill retention could be evaluated over short or long periods. For example, Project READS seeks to mitigate summer reading loss by sending books to students' homes. This program could be evaluated using EDU STAR if baseline assessments were conducted for treatment and control groups at the end of the school year with a follow-up assessment in the fall. The evaluation costs of such programs would dramatically decline where the EDU STAR platform was introduced.

While such longer-term studies would not be as rapid or low cost as real-time evaluations of digital learning technologies, they would be far simpler to conduct than is currently possible. In essence, by automating the preintervention steps (layers 1 and 2) and the postintervention steps (layers 4 and 5) in the evaluation schematic above, the intervention layer can ultimately be handled quite flexibly.

ii. The Reporting System

Once EDU STAR has collected sufficient data on the effectiveness of software applications against specific CCSS, it can rank these products relative to the baseline and relative to each other. EDU STAR could rate learning technologies between one and five stars depending on their comparative effectiveness against the placebo, where relative success is measured in standard deviations over the baseline performance. Software packages that include multiple programs covering

FIGURE 4.

Example of EDU STAR Reporting with Common Core State Standards

(1) Standards Page

Standard	Description	EDU STAR Reports
L.3.1	General: Demonstrate command of the conventions of standard English grammar and usage when writing or speaking. (Click for detailed description.)	Click here.
L.3.2	Demonstrate command of the conventions of standard English capitalization, punctuation, and spelling when writing. (Click for detailed description.)	Click here.
. . .	. . .	. . .

(2) Reports Page

Standard L.3.2

Demonstrate command of the conventions of standard English capitalization, punctuation, and spelling when writing. (Click for detailed description.)

List of EDU STAR Reports for Standard L.3.2

Product Name	Rating	Evaluation	Company
ABCD Reading	*****	Detailed Analysis	Readers Company (Click here.)
XYZ Reading	***	Detailed Analysis	Education Inc. (Click here.)
12345 Reading	**	Detailed Analysis	Education Solutions (Click here.)

(3) Detailed Analysis Page

Product: XYZ Reading
XYZ Reading has been tested using randomized trials involving more than 4,200 students across thirty states. Based on these results, users of the software improved their performance on assessments of standard L.3.1 by an average of 0.15 standard deviations. There were no statistically significant differences by gender, race, or initial skill level, however. . . .

multiple CCSS would be rated within each standard to avoid commensurability challenges. Rated technologies that undergo further innovations would continue to be evaluated to provide useful feedback to producers and customers.

The results would be available on a new, public website. Users will review evaluated results, organizing lists of learning tools according to specific CCSS. The star rating system will provide simplicity and transparency for the nonexpert, modeled after the intuitive systems that *Consumer Reports* uses today. At the same time, clicking through the top-level scores would allow users, including teachers and procurement officers, to access information about evaluation methods and results that is more detailed. This information would explain how many students have used the software, how it was tested, teacher- and student-generated ratings, and whether there was evidence of heterogeneous treatment effects (i.e., the tool has variable impact depending on observable student characteristics, such as their baseline skill level). Figure 4 provides a rough mock-up of how the data structure could work. The reported information could also be integrated with other relevant websites, such as the CCSS initiative (http://www.corestandards.org/).

EDU STAR's reporting system would be designed to maximize the credibility of the organization as an independent and fair evaluator. First, all evaluations would be registered publically on the website, regardless of their outcome, so that sellers cannot cherry-pick the results. Second, while EDU STAR may eventually accept user fees from technology producers, it would not accept any other grants or investments from organizations with a financial stake in its evaluations. Similar to *Consumer Reports*, EDU STAR would not allow commercial advertisements on its website, nor would it accept free samples. Third, employees of

EDU STAR would adhere to strict rules regarding participation in industry-funded conferences or sponsored travel. Along with a transparent evaluation methodology, these steps will help establish the credibility of EDU STAR.

iii. Building the Test Bed

For EDU STAR to generate useful data, it must be able to facilitate trials with sufficiently large numbers of students. In particular, to develop comparative effectiveness insights between two different applications with similar performance characteristics, larger sample sizes will be required.

At launch, EDU STAR should ideally have several schools and a significant number of students signed up to participate. Fortunately, there are several ongoing efforts within schools to test new technologies (including Gates Foundation–sponsored initiatives) and draw together schools interested in innovation (Digital Promise's League of Innovative Schools). These efforts can support the creation of initial test beds for EDU STAR.

Note that participation in the evaluation platform need only appeal to a small set of schools. As a rule of thumb, the inclusion of one large school system in a major city would meet the sample size and power requirements necessary to generate useful results. We envision EDU STAR attracting interest among public, charter, and private schools; many of these schools are eager to engage with leading-edge learning tools. Nonetheless, if necessary, incentives can be provided to increase participation early on. A natural incentive would be price discounts for evaluated educational technologies. This subsidy could be provided by the innovator (who benefits from using the test bed) or possibly by government, private, or nonprofit funders of EDU STAR. (See also the discussion of potential funders and partners in chapter 3.E.i.)

The EDU STAR test bed can work flexibly around differing school schedules. Schools may dedicate part of their existing computer lab time to working with these innovative tools. Alternatively, participating schools could build explicit computer time into existing class schedules or schedule special computer time after school hours. Some schools may choose to start the day as School of One does, with an online component. As an "always ready" Internet-based system, EDU STAR can queue software to be evaluated and allocate the evaluations seamlessly whenever schools schedule computer time for their students. It can also flexibly add or remove schools as they enter or leave the test bed.

iv. Linking to Innovators and Entrepreneurs

We expect EDU STAR to be immediately popular with innovators and entrepreneurs. In particular, individuals and companies with good products but little market presence will be eager to participate in a test bed that demonstrates effectiveness, which will give them an enormous marketing advantage.

EDU STAR's functions vis-à-vis innovators and entrepreneurs include the following:

- Initial outreach, so that existing companies and entrepreneurs learn of the platform's existence. This outreach can include communications with existing educational software companies, teachers, entrepreneurial incubators, education-school students and faculty, and business students, as well as angel investor and venture capital networks.
- Screening, once the system is running, to ensure that any software distributed on the test bed is of sufficient quality.
- Beta-testing, allowing prepublic evaluation of software, to provide innovators with feedback on their products and indicate whether they succeed.
- Formal evaluation of publically available products.

Similar to the creation of the appstore platforms that have sparked tens of thousands of entrepreneurs to create new products, we expect EDU STAR to have a dramatic impact on the innovation ecosystem in education.

E. LAUNCHING EDU STAR

While it is challenging to estimate precise costs for the EDU STAR platform, we believe that (a) total costs will be low and (b) there are several natural sponsors. In terms of financing, one can assess potential costs by considering each of EDU STAR's core functions.

- In linking to innovators, staff is necessary in the short run to market the platform. Staff is also necessary in the longer run to screen software proposed by innovators and to communicate with innovators over beta-testing and formal evaluations.
- In linking to schools, staff is necessary in the short run to find participants for the test bed. As demand for the platform increases, the initial test bed, depending on its size, may also need to expand. Ongoing costs here will largely involve communications with participating schools.
- In building the test bed, up-front costs include developing the cloud-computing servers and software to run the evaluations in an automated fashion (the layers in the evaluation schematic). The programming task here seems straightforward; the evaluation methodology and algorithms are well known, and very similar systems are already implemented by numerous corporations. Once this system is created, maintenance costs will remain, and the system will be best served with some continual development.
- In reporting, the up-front costs involve disseminating the evaluation results as described above. This web-based system should be low cost to develop and maintain.

A (short) technical study needs to be performed to develop an accurate forecast of costs. We believe that the major initial expenditures involve a small technical staff (outsourcing programming), a small marketing and communications staff, and some very low-cost investments in computer servers. We therefore anticipate that EDU STAR could be successfully developed and launched with five full-time employees and roughly a $5 million budget. Furthermore, EDU STAR would include an advisory board of ten members including leading teachers, entrepreneurs, IT professionals, educational policy experts, education researchers, and school officials. As detailed below, many of the other necessary inputs to start-up may be available as in-kind donations. The steps in launching EDU STAR are summarized in Box 2.

i. Funders

There are several natural and alternative sponsors for the creation of EDU STAR. We propose that EDU STAR be launched as a third-party 501(c)3 organization, initially seeded by foundation and government grants. For example, the Gates Foundation, the MacArthur Foundation, and the Broad Foundation, among others, already have deep interests and make substantial investments in education technology. Furthermore, private corporations may also be willing to provide technical leadership and services. For example, for-profit companies like Google or Amazon may be willing to provide programming expertise in developing the rapid evaluation platform, and companies like Microsoft, Amazon, Apple, and others may be willing to provide a mix of computer hardware or cloud-computing services pro bono.

There are promising opportunities for government funding of EDU STAR as well, most notably the Investing in Innovation (i3) fund in the U.S. Department of Education, a $150 million program. EDU STAR could naturally follow a sequence of i3's "development grants," "validation grants," and "scale-up grants" as needed, which had maximum awards of $3 million, $15 million, and $25 million, respectively, in FY2011.

BOX 2.

Launching EDU STAR

- **EDU STAR can be launched as a 501(c)3 non-profit organization with $5 million in funding, drawing on a consortium of foundations, private sector partners, or a U.S. Department of Education Investing in Innovation (i3) grant. EDU STAR will rent its own office space or be housed in the facilities of one of the consortium members.**
- **At launch, EDU STAR would have five full-time employees: executive director, director of research, chief technical officer, director of development, and office manager.**
    - **The executive director would launch and run the overall organization, create an advisory board, and build relationships with stakeholders. The initial executive director would ideally start from an existing technology-focused educational nonprofit or foundation, raising money and launching EDU STAR from within an existing organization.**
    - **The director of research would develop detailed evaluation and reporting design principles, working with outside academics, business people, government officials, and policy experts, and convening an institutional review board (IRB) to oversee evaluation procedures.**
    - **The chief technical officer would coordinate development of the technology infrastructure for EDU STAR, working with private sector technical experts and school IT officials, and overseeing vendors in building the Internet-based evaluation and reporting infrastructure.**
    - **The director of development would conduct outreach to (a) school districts and (b) innovator communities in building the test bed. The director of development also would manage finances and fund-raising.**
    - **The office manager would handle administrative functions of EDU STAR.**
- **The advisory board of ten members would include leading teachers, entrepreneurs, IT professionals, educational policy experts, education researchers, and school officials.**
- **The EDU STAR platform initially would be limited to instructional software tied to specific CCSS and appropriate for a typical class period. One year after the platform's launch, EDU STAR staff and board would consider including expanded educational interventions.**
- **User fees, initially absent, would grow progressively to eventually cover all of the operations for EDU STAR. User fees will be assessed per technology application with a sliding scale featuring low fees for small businesses.**

Building the EDU STAR platform would also be a natural target for Advanced Research Projects Agency–Education (ARPA-ED), which was proposed but not funded for FY2012 with the purpose of funding innovative private sector, university, and nonprofit projects to generate dramatic breakthroughs in education. The proposed annual budget for ARPA-ED was $90 million.

In the longer run, as the system becomes established, and especially with its innately low-cost operations, it can move to a self-sustaining funding model. In particular, technology producers could pay user fees to EDU STAR in proportion with their size (similar to FDA user fees conceptually, but with much lower fees in practice). Since EDU STAR would be conducting rapid randomized controlled trials, similar to what Google and Amazon do hundreds of times per day, we do not anticipate significant additional costs after initial launch.

ii. Partners

Institutionally, EDU STAR can form useful partnerships with existing nonprofits, who have interrelated initiatives.

- Digital Promise, with its League of Innovative Schools, provides a set of existing school systems committed to innovation. These schools are likely willing participants in the test bed; by participating with the League of Innovative Schools, EDU STAR may simplify the test bed's creation.
- The Susan Crown Exchange, working with Common Sense Media, is developing a ratings system for digital learning technologies. While their platform neither focuses (yet) on rigorous evaluation metrics nor considers the CCSS, it could become an outlet and partner for rigorous evaluations.
- CFY.org, which collects and delivers digital learning content in the after-school (i.e., at-home) market, could provide additional test beds and a reporting outlet for reaching parents in addition to schools.
- The Common Core State Standards Initiative could partner with EDU STAR, promoting its mission and reporting among state schools systems, given the natural pairing between EDU STAR and the Common Core State Standards Initiative itself.

Finally, specific school systems could play leading roles. Schools that already use testing and personalized learning, like the School of One in New York City, already have infrastructure in place to provide a test bed for new technologies. To the extent that evaluated software is part of larger modules or courses, EDU STAR results could also support initiatives like Carnegie Mellon's evaluation of open course content.

F. TIMELINE

By using the Internet platform, the technical pieces for the evaluation system can be developed independently from specific participating schools or innovators. Hence (a) the evaluation system, (b) the school-facing test bed commitments, and (c) the innovator-facing outreach activities can all occur in parallel in the initial development phase. The reporting format could be developed once the evaluation system is up and running, but ideally the reporting format will be developed in parallel so that its design details feed upstream into the design of the evaluation system. Overall, while a (short) technical study needs to be performed, we expect that the EDU STAR platform could launch within eighteen months of initial funding.

G. OTHER STEPS

Several other steps are not necessary for EDU STAR, but would further increase the power of the evaluation and reporting platform.

i. Linking to Existing Databases

First, anonymous evaluation results could be linked to other (anonymous) demographic information, such as interventions that the student has previously participated in and prior academic performance. Many schools have databases that could be linked to the EDU STAR results. Tapping this information would lead to a more informed understanding of heterogeneous treatment effects and lay the foundation for personalized learning, an analogue to the personalized medicine revolution in health.

ii. Data Standards

Such database linkages would be easier with database standards, which will more generally allow better interoperability of software solutions. For example, with common data-sharing protocols, learning software from different companies can more seamlessly interact—facilitating innovation on individual learning modules. In the absence of standardized protocols, technology companies will tend to engage in costly, noninteroperable learning systems. Teachers and schools will find it cumbersome to use software with unique protocols. Furthermore, different learning software would need to start over with each student, rather than building on existing knowledge.[14]

To be clear, we would not suggest that such data standardization be done by EDU STAR, and such data standards are not necessary for EDU STAR to provide evaluation and reporting. Nonetheless, such data standards could enhance EDU STAR's mission while improving the utility of learning technologies more generally.

iii. Prizes

Prize competitions can target specific training objectives, following the detailed standards of the CCSS and with prize-winning criteria defined by successful evaluation. Prize competitions would be an exciting way to build community support and encourage the creation of new software, and conceivably could draw on entries from student teams.

Chapter 4: Addressing Potential Challenges and Concerns

There are several potential challenges and concerns that might impact the development and implementation of EDU STAR. These include regulatory matters around privacy and research, stakeholder interests, and fidelity of implementation issues. We consider these issues next.

A. PRIVACY AND RESEARCH REGULATIONS

Regulations from both the Family Education Rights and Privacy Act (FERPA) and IRBs intersect with the EDU STAR platform. FERPA defines when parental consent is and is not required for release of student data, and allows circumscribed data use by organizations for evaluation purposes. By emphasizing the use of anonymous data (i.e., by not collecting names) and using secure data storage systems, these regulations can traditionally be met. Regarding research regulations, which work to ensure that research is done in an ethical manner, we recommend the use of an IRB to advise and approve the design and implementation of EDU STAR.

Regulatory requirements have long been met for individual assessments of educational technologies, and New York City's School of One provides a template for scaling these systems. The key for EDU STAR will be to avoid requiring separate consent for every independent application or evaluation exercise, which would of course substantially raise implementation costs. Currently, Digital Promise is working on blanket FERPA and IRB consent for its League of Innovative Schools, providing a template for EDU STAR and a potential test bed avenue within that community of school systems.

B. STAKEHOLDER INTERESTS

It will be important for EDU STAR to communicate its mission accurately and transparently to stakeholders so that they can assess the potential benefits and choose to participate in (or at least not obstruct) development of the EDU STAR platform. The social benefits of any low-cost system that can improve elements of K–12 education, enhancing individual opportunity and the success of the economy at large, seem clear. The benefits of reducing entry barriers for would-be innovators and entrepreneurs will naturally align interests on the supply side, as discussed above. Next, we focus on the benefits for teachers and parents.

Some observers believe that teachers and technology are fundamentally opposed, assuming they are substitutes in delivering education. We believe that this perspective misses the first-order point that learning tools augment rather than replace teaching time and align naturally with teachers' desire to educate their students as effectively as possible.

By seeing these tools as "teacher's assistants," these tools become ways to personalize learning and temporarily engage subsets of students, freeing the teacher to focus on students with particular needs and further enhancing teacher and student success.

Digital learning tools are akin to textbooks, workbooks, or homework exercises—tools in the teacher's toolkit, which the teacher deploys holistically to achieve instructional goals. Furthermore, EDU STAR will only cover a portion of the teacher's toolkit: the platform will work best when applied to discrete and measurable skills, like counting and reading, and only those in the CCSS, and will be less appropriate for other skills, such as expository writing, creative work, presentation skills, and so on. Thus, we view this platform as a complement to other important teaching dimensions. Moreover, by seeing these tools as "teacher's assistants," these tools become ways to personalize learning and temporarily engage subsets of students, freeing the teacher to focus on students with particular needs and further enhancing teacher and student success. Finally, we predict that many of the best product ideas will come from teachers themselves. With expert teachers as the likely source of highly effective learning ideas, EDU STAR will include a program to pair star programmers with expert teachers (e.g., teachers who win state and national Teacher of the Year awards) to facilitate the codevelopment of new tools.

Parents will naturally align with EDU STAR goals as they seek to improve educational outcomes for their own children. Many parents are also likely to adopt proven digitally delivered learning tools at home, through home computers and mobile devices. When technologies are proven and the results are provided to parents, parents will also likely help guide school systems to adopt best practices. However, parents may also have concerns about EDU STAR, especially as it relates to privacy and concerns about overtesting students. First, it is essential that system data be anonymous to protect student identity, as discussed above. Related, it is also important to emphasize to parents that this system evaluates tools, not students. The results of EDU STAR evaluations should not be linked to course grades. It will also be important that placebo groups will not involve students interacting with products that do not provide any benefits, but rather involve existing lessons plans and technologies.

More generally, schools that join the platform will need to consult parents early on and provide the requisite detail to determine whether joining the test bed is right for their school community. A recommended approach might be to launch a pilot program that can be assessed by all stakeholders before a school formally joins EDU STAR.

While some incentives may be required to increase participation, note again that EDU STAR needs the participation of only a small number of schools from among the vast array of schools—public, charter, and private—in the U.S. K–12 education landscape. Just as a small number of research hospitals lead the way to better health for a much larger population, the evaluation and reporting functions of EDU STAR can benefit school systems at large even though the test bed itself can be very small.

C. FIDELITY OF IMPLEMENTATION

Another important concern is around fidelity of implementation. Will tools be used as intended and hence assessed accurately? How will the interventions be managed effectively? We believe that the design of the EDU STAR system, in schools that choose to participate, will avoid fidelity of implementation challenges. First, the system randomizes the students across tools, automates a common assessment protocol, and delivers the tools to be evaluated. The teacher will have to make sure that students are not helping each other with their assessments to prevent contamination of the data, but this would be a requirement under most conditions anyway. More generally, the students might not take the assessment seriously and type in random answers to questions. However, while this challenge is endemic to all kinds of testing and evaluation, there is a built-in remedy with EDU STAR: we imagine that the highest-rated educational tools will be those that engage students through fun and absorbing exercises. Thus, by design EDU STAR evaluations will implicitly favor technologies that keep students engaged in the material, incorporating and reducing fidelity of implementation challenges that can be construed as flaws in the tool design itself. Good product designers do not blame the user—they design products that the user finds engaging and transparent to use.

Chapter 5: Conclusions

By taking America's strength in innovation and applying it to our K–12 sector, we can create an innovation ecosystem that produces effective and continuously improving learning technologies. To accomplish this goal, we argue that a simple and low-cost institutional innovation is required. Evaluating and reporting on learning tools that work against CCSS will empower teachers to make smart decisions in technology adoption and provide incentives for innovators to develop products for the K–12 market. If EDU STAR is successful, we will see a simple and relatively inexpensive institutional innovation reshape the learning technology market, create a rush of innovation, and ultimately increase the quality of U.S. primary and secondary education.

Authors

Aaron Chatterji

Associate Professor
Fuqua School of Business, Duke University

Aaron Chatterji, Ph.D. is an Associate Professor at Duke University's Fuqua School of Business. He previously served as a Senior Economist at the White House Council of Economic Advisers where he worked on a wide range of policies relating to entrepreneurship, innovation, infrastructure and economic growth. Chatterji's research and teaching investigate some of the most important forces shaping our global economy and society: entrepreneurship, innovation, and the expanding social mission of business. He was awarded an inaugural Junior Faculty Fellowship from the Kauffman Foundation to recognize his work as a leading scholar in entrepreneurship. He also received the Rising Star award from the Aspen Institute for his work on business and public policy.

His research has been published in leading academic journals and been cited by The New York Times, CNN, The Wall Street Journal, and The Economist. He has authored several op-ed pieces in leading newspapers, appeared on TV and radio, and was recently profiled in The Financial Times and Fortune. Chatterji is a term member of the Council on Foreign Relations and previously worked as a financial analyst at Goldman Sachs. He received his Ph.D. from the Haas School of Business at the University of California at Berkeley and his B.A. in Economics from Cornell University.

Benjamin Jones

Associate Professor of Management & Strategy
Kellogg School of Management, Northwestern University

Benjamin F. Jones is an Associate Professor at the Kellogg School of Management, where he directs the Kellogg Innovation and Entrepreneurship Initiative, and is a faculty research fellow of the National Bureau of Economic Research. An economist by training, his research focuses largely on innovation and creativity, with recent work investigating the role of teamwork in innovation and the relationship between age and invention. Professor Jones also studies global economic development, including the roles of education, climate, and national leadership in explaining the wealth and poverty of nations. His research has appeared in journals such as Science and the Quarterly Journal of Economics and has been profiled in media outlets such as the Wall Street Journal, the Economist, and The New Yorker.

Professor Jones served in 2010-2011 as the senior economist for macroeconomics for the White House Council of Economic Advisers and earlier served in the U.S. Department of the Treasury. In 2011, he was awarded the Stanley Reiter Best Paper Award for the best academic article written by a Kellogg faculty member in the prior four years.

Endnotes

1. The 0.2 percent estimate comes from the President's Council of Advisors on Science and Technology (2010). This figure is somewhat higher than other estimates.
2. Google and other Internet companies have demonstrated that evaluations can be especially rapid and low-cost when variations in software are being evaluated in real time across a large-scale test bed. See further discussion below.
3. Nor is it surprising that computer hardware alone, in the absence of a high-quality instructional content layer, have little impact on student performance (Angrist and Lavy 2002; Goolsbee and Guryan 2006).
4. However, the variation in expenditure patterns per student and other idiosyncratic features of these large school systems suggests challenges for educational technology companies in building and marketing their products even across these big districts.
5. Berger and Stevenson (2007) report three additional companies, including Scholastic, with annual revenues in the high hundreds of millions of dollars—approximately a dozen companies in the $100 million to $250 million range, a small group in the $25 million to $100 million range, and many smaller companies.
6. Education Division, Software & Information Industry Association (last visited March 14, 2011), http://www.siia.net/index.php?option=com_content&view=article&id=141&Itemid=169. It is not clear precisely how many are involved in the software space for K-12.
7. Market participants suggest numbers in the $50 million to $250 million range for venture backing in educational technology; the reported estimates comes from a National Venture Capital Association's Spotlight on Education study (Mendell 2010). Estimates for other industries come from PriceWaterhouse Coopers (2010).
8. Berger and Stevenson (2007) note that there is more money for building schools than for building tools because decentralized governance creates room for new schools but no integrated market for tools.
9. One observer also notes, however, that the recent purchase of Wireless Generation by the News Corporation has created excitement and a perception of more numerous exit options.
10. There are of course other challenges—for example, it may simply be difficult to design effective educational technologies. Also, successful educational software needs to be highly engaging for the student and consistent with the incentives of teachers and school administrators who are asked to deploy these technologies. However, should policy create the right environment for an educational technology market, then businesses, nonprofits, entrepreneurs, and inventors (including teachers or even students themselves) will have a pathway for innovation. When the rules are right, we can count on decentralized actors to bring remarkable new ideas, as they do in many other sectors of the economy.
11. Important points of contact can include state and local procurement officers who oversee funding streams, academic consultants who advise districts, key school board members, directors of IT, and principals and teachers in individual schools.
12. Educational technology companies often complain that, even when their products are purchased, inadequate training in their use limits their utility. It is less clear, however, whether this problem is in the design of products and insufficient support from the sellers, or whether there is an issue on the user side.
13. Nor is it the What Works Clearinghouse, a reporting (but not evaluation) organization in the Department of Education, which collects the results of one-off research studies. While very well motivated, a challenge for the What Works Clearinghouse has simply been the thin pipeline of rigorous evaluations, which are traditionally slow and expensive to perform, leaving the clearinghouse with little to report on. The EDU STAR system aims to solve this root challenge by providing rigorous evaluation at far lower cost for learning tools.
14. Interoperability is an additional reason that large, established players in the marketplace currently have an advantage, in that they can offer schools a suite of products that work in tandem. At the same time, this issue forces schools into large commitments to particular sellers, an investment that can scare schools away in the first place or, if schools commit, trap buyers into paying high prices later rather than paying the high costs of switching to a new system. The alternative model, envisioned here, is one with modular software that interacts in a standardized way, so that buyers are not committed to the product of one company. Like the Internet or the iPad, standardized protocols allow entrepreneurial entry, where users can fluidly draw on a mix of products across firms.
15. Even large incumbent firms can benefit from lower marketing costs and greater total market size.
16. For example, pharmaceutical companies have mechanisms for paying participants in clinical trials. While we remain skeptical that such incentives will be needed for the EDU STAR platform, small incentive payments for schools, or classrooms within schools, can be considered if needed.

References

Angrist, Josh, and Victor Lavy. 2002. "New Evidence on Classroom Computers and Pupil Learning." *Economic Journal* 112.

Barrow, Lisa, Lisa Markman, and Cecilia E. Rouse. 2009. "Technology's Edge: The Educational Benefits of Computer-Aided Instruction." *American Economic Journal Economic Policy* 1 (1).

Berger, Larry, and David Stevenson. 2007. "K–12 Entrepreneurship: Slow Entry, Distant Exit." Wireless Generation mimeo. Accessed at http://www.aei.org/files/2007/10/25/20071024_BergerStevenson.pdf.

Bureau of Economic Analysis (BEA). 2012. Industry Economic Accounts 1998-2011. Accessed at http://www.bea.gov/industry/iedguide.htm#gdpia_ad_it.

Council of Economic Advisers. 2011, September. "Unleashing the Potential of Education Technology." Accessed at http://www.whitehouse.gov/sites/default/files/unleashing_the_potential_of_educational_technology.pdf.

EPE Research Center. 2010. "Graduation Rates in the Nation's 50 Largest School Districts." Accessed at http://www.edweek.org/media/34arrow-c1.pdf.

Goldin, Claudia, and Lawrence F. Katz. 2007. "Long-Run Changes in the Wage Structure: Narrowing, Widening, Polarizing." Brookings Papers on Economic Activity 2:135–65. Brookings Institution, Washington, DC.

Goolsbee, Austan, and Jonathan Guryan. 2006. "The Impact of Internet Subsidies in Public Schools." *Review of Economics and Statistics* 88.

Manzi, Jim. 2012. *Uncontrolled: The Surprising Payoff of Trial-and-Error for Business, Politics, and Society*. New York: Basic Books.

Mendell, Emily. 2010. "NVCA Spotlight on Education" Online posting. 15 November 2010. Accessed at http://nvcaccess.nvca.org/index.php/topics/commentary/151-nvca-spotlight-on-education.pdf.

National Science Foundation (NSF). 2011. "Research and Development in Industry: 2006–07." Accessed at http://www.nsf.gov/statistics/nsf11301/pdf/nsf11301.pdf.

National Science Foundation (NSF). 2012. "Research and Development: National Trends and International Comparisons." In *Science and Engineering Indicators 2012* (Chapter 4). Accessed at http://www.nsf.gov/statistics/seind12/pdf/c04.pdf.

NeXtup. 2010. "NeXt Knowledge Factbook 2010." Next Up! Research. Accessed at http://www.kellogg.northwestern.edu/faculty/jones-ben/htm/NextKnowledgeFactbook2010.pdf.

Organisation for Economic Co-operation and Development (OECD). 2009. "Education at a Glance 2009." Accessed at http://www.oecd.org/education/highereducationandadultlearning/43636332.pdf.

Organisation for Economic Co-operation and Development (OECD). 2010. "PISA 2009 Results, Figure 1." Accessed at http://www.oecd.org/dataoecd/54/12/46643496.pdf.

Organisation for Economic Co-operation and Development (OECD). 2011. "Education at a Glance, Chapter B: Financial and Human Resources." Accessed at http://www.oecd.org/education/highereducationandadultlearning/48630868.pdf.

President's Council of Advisors on Science and Technology. 2010. "Prepare and Inspire: K–12 Science, Technology, Engineering, and Math (STEM) Education for America's Future." Accessed at http://www.whitehouse.gov/sites/default/files/microsites/ostp/pcast-stem-ed-final.pdf.

PriceWaterhouse Coopers. 2010. "National Venture Capital Association MoneyTree Report." Accessed at http://www.nvca.org/index.php?option=com_docman&task=cat_view&gid=57&Itemid=317.

Rouse, Cecilia E., and Alan B. Krueger. 2004. "Putting Computerized Instruction to the Test: a Randomized Evaluation of a 'Scientifically Based' Reading Program." *Economics of Education Review* 23 (4, August).

Stillwell, Robert, Jennifer Sable, and Chris Plotts. 2011, May. "Public School Graduates and Dropouts from the Common Core of Data: School Year 2008–09." National Center for Education Statistics Paper 2011-312.