Video games: A route to large-scale STEM education?

1 author:

Merrilea Joyce Mayo
Mayo Enterprises, LLC

79 PUBLICATIONS   **3,448** CITATIONS

implementation of their science assessment systems (*21*). In a report commissioned by that project, Quellmalz and Moody (*22*) proposed strategies for states to form collaboratives and use technology to create multilevel science assessment systems. With the goal of helping schools and students meet the NCLB goals, states are seeing classroom-based, instructional uses of assessment as a powerful tool for driving student achievement. Such assessment is distinguished from interim assessments administered periodically on a larger scale that are intended to describe the status of student performance after instruction (*23*).

A key feature in creating a balanced multilevel system is the use of common design specifications that can operate across classroom, district, state, and national levels (*22*). To enable implementation, on-line authoring systems are being developed that can assist in creating such common specifications, in streamlining test design, and in reducing development costs (*24*). Online design systems can also support adaptations of assessments to offer accommodations for special populations while preserving the linkages between targeted standards and designs of the tasks for eliciting evidence of achievement.

Conclusion

Technology helps us do many conventional things in the world of testing and assessment better and faster, and it holds the key to transforming current assessment practice for multiple purposes and at multiple levels ranging from the classroom to state, national, and international levels. We are not there yet, and although many obstacles remain to their widespread use, the next generation of technology-enabled assessments is under development with several promising cases of design, implementation, and use. Such demonstrations provide a vision of the possible and can help move education toward the design and adoption of more integrated and effective learning-centered assessment tools and systems.

References

1. National Center on Education and the Economy, "Tough choices or tough times: Report of the new commission on the skills of the American workforce" (Jossey Bass, Washington, DC, 2006).
2. National Research Council, *Rising Above the Gathering Storm: Energizing and Employing America for a Brighter Economic Future* (National Academies Press, Washington, DC, 2006).
3. R. E. Bennett, *Educ. Policy Anal. Arch.* **9**, 5 (2001).
4. R. E. Bennett, "Technology for large-scale assessment" (ETS Report No. RM-08-10, Educational Testing Service, Princeton, NJ, 2008).
5. M. Koomen, paper presented at the Annual Meeting of the American Educational Research Association, San Francisco, CA, April 2006.
6. B. Sandene et al., "Online assessment in mathematics and writing: Reports from the NAEP technology-based assessment project" (NCES 2005-457, U.S. Department of Education, National Center for Educational Statistics, U.S. Government Printing Office. Washington, DC, 2005).
7. T. K. Landauer, D. Laham, P. Foltz, *Assess. Educ.* **10**, 295 (2003).
8. S. Klein, in *Probability and Statistics: Essays in Honor of David A. Freedman,* D. Nolan, T. Speed, Eds. (Institute of Mathematical Statistics, Beachwood, OH, 2008), vol. 2, pp. 76–89.
9. T. Vendlinski, R. Stevens, *J. Technol. Learn. Assessment* **1**, 3 (2002).
10. R. E. Bennett, H. Persky, A. Weiss, F. Jenkins, "Problem solving in technology rich environments: A report from the NAEP technology-based assessment project" (NCES 2007-466, U.S. Department of Education, National Center for Educational Statistics, U.S. Government Printing Office, Washington, DC, 2007).
11. P. Black, D. Wiliam, *Inside the Black Box: Raising Standards Through Classroom Assessment* (King's College, London, 1998).
12. D. Wiliam, in *Second Handbook of Mathematics Teaching and Learning,* F. K. Lester Jr., Ed. (Information Age Publishing, Greenwich, CT, 2007), pp. 1051–1098.
13. P. Black, C. Harrison, C. Lee, B. Marshall, D. Wiliam, *Phi Delta Kappan* **86**, 8 (2004).
14. J. Brown, S. Hinze, J. W. Pellegrino, in *21st Century Education,* T. Good, Ed. (Sage, Thousand Oaks, CA, 2008), vol. 2, chap. 77, pp. 245–255.
15. J. Minstrell, P. Kraus, in *How Students Learn: History, Mathematics, and Science in the Classroom,* J. Bransford, S. Donovan, Eds. (National Academies Press, Washington DC, 2005), chap. 11.
16. A. Thissen-Roe, E. B. Hunt, J. Minstrell, *Behav. Res. Meth. Instrum.* **36**, 234 (2004).
17. M. Feng, N. Heffernan, K. Koedinger, in *Proceedings of the Eighth International Conference on Intelligent Tutoring Systems,* M. Ikeda, K. D. Ashley, T. W. Chan, Eds. (Springer-Verlag, Berlin, 2006), pp. 31–40.
18. E. S. Quellmalz, G. Haertel, "Technology supports for state science assessment systems" (National Research Council, Washington, DC, 2004).
19. E. S. Quellmalz et al., in *Assessing Science Learning: Perspectives from Research and Practice,* J. Coffey, R. Douglas, C. Stearns, Eds. (National Science Teachers Association Press, Washington, DC, 2008), chap. 10.
20. J. Pellegrino, N. Chudowsky, R. Glaser, Eds., *Knowing What Students Know: The Science and Design of Educational Assessment.* (National Academies Press, Washington, DC, 2001)
21. M. R. Wilson, M. W. Bertenthal, Eds., *Systems for State Science Assessment* (National Academies Press, Washington, DC, 2005).
22. E. S. Quellmalz, M. Moody, "Models for multi-level state science assessment systems" (National Research Council, Washington, DC, 2004).
23. M. Perie, S. Marion, B. Gong, "A framework for considering interim assessments" (National Center for the Improvement of Education Assessment, Dover, NH, 2007).
24. R. Mislevy, G. D. Haertel, *Educ. Meas.* **25**, 6 (2006).

PERSPECTIVE

Video Games: A Route to Large-Scale STEM Education?

Merrilea J. Mayo

Video games have enormous mass appeal, reaching audiences in the hundreds of thousands to millions. They also embed many pedagogical practices known to be effective in other environments. This article reviews the sparse but encouraging data on learning outcomes for video games in science, technology, engineering, and math (STEM) disciplines, then reviews the infrastructural obstacles to wider adoption of this new medium.

In the 2000-to-2005 time frame, ~450,000 students graduated annually in the United States with a bachelor's degree in STEM (*1*). These numbers pale in comparison to the reach of a single computer video game (Figs. 1 and 2). *World of Warcraft* (*2*), a fantasy game, has over 10 million current subscribers, with ~2.5 million in North America (*3*). *Food Force* (*4*), the U.N.-produced game on the mechanics of food aid distribution, saw 1 million players in its first 6 weeks and 4 million players in its first year (*5*). Additionally, in the realm of K-to-12 science and math education, the virtual world *Whyville* (*6*), with its game-based activities, now sports 4 million subscribers (90% North American), with the dominant demographic being 8- to 14-year-old girls (*7*, *8*). Although traditional education institutions pride themselves on educating citizens, they do so at a relatively small scale compared with the media now available. Is it possible to greatly expand the reach of STEM education with the use of video games as the medium? And to what level of effectiveness?

At first, the idea of using video games to teach science and engineering seems laughable. However, sophisticated video game content already exists in topics ranging from immunology (*9*) (Fig. 3) to numerical methods (*10*, *11*). The examples in Table 1 suggest that video games can yield a 7 to 40% positive learning increase over a lecture program. What's more, there may be additional benefits to poor learners: One variant of the *River City* ecology game (*12*) diminished the learning gap between D and B students to the point where nearly all students were performing at the B-student level (*13*).

Learning outcomes are by no means uniformly positive. Results from review studies (*14*, *15*) make

Ewing Marion Kauffman Foundation, 4801 Rockhill Road, Kansas City, MO 64110 USA. E-mail: mmayo@kauffman.org

Education & Technology

it clear that there are both well-designed games and poorly designed ones. Where learning benefits appear, they are attributed to effective pedagogical practices embedded in the game design (14–17). Of course, many of these same practices can also be applied to classroom, Web, or other forms of instruction with similar benefits, an approach known as game-informed learning (18).

Unlike lectures, games can be adapted to the pace of the user. Games also simultaneously present information in multiple visual and auditory modes, which capitalizes on different learning styles. J. P. Gee (16) identifies the former as the "just-in-time principle" and the latter as the "multimodal principle" in his book on video game–based learning (16), reviewed in (19). Games are also particularly adept at dosing information delivery. Complex tasks are presented first as a small core experience that is practiced multiple times before being progressively extended into a longer, more complex sequence. The superior efficiency of this approach (known as concurrent chaining) has been compared with whole-task learning in (20). Gee (16) describes this kind of task structuring through his "incremental principle," "concentrated sample principle," and "bottom-up basic-skills principle."

Games are also useful for reinforcing information acquisition. The rich environment of objects and activities within games gives information "situated meaning": the other contextual elements support the information being conveyed. Social surroundings can also reinforce content. Well-constructed social interactions around societal goals within the game will drive learner engagement and achievement, as has been studied in depth by S. Barab *et al.* in their *Quest Atlantis* project (21, 22). Content is further reinforced through continuous, immediate feedback: Almost every keystroke yields a response from the game. In contrast, students in a typical classroom get to ask 0.11 questions per hour (23). And, finally, a steady stream of positive rewards accompanies a game's rapid feedback. Players accumulate points, levels, titles, or magic swords with some visible progress for even the tiniest successes. These rewards contribute to greater self-confidence/self-efficacy. Greater self-efficacy, in turn, translates to greater

persistence and thus a higher level of accomplishment (24).

Learner control over navigation through tasks and activities is a surprisingly important feature of effective learning games. The metastudy by J. J. Vogel *et al.* (15) found learner control/ autonomy to be one of the few easily identified predictors of enhanced learning outcomes,

Fig. 1. Comparison of online game subscriptions (3, 7) to U.S. bachelor's degrees awarded across all STEM disciplines (1) as well as in just the engineering disciplines (1). Games having more than 1 million subscribers are shown.

Fig. 2. Comparison of online game subscriptions (3) to U.S. bachelor's degrees awarded across all STEM disciplines (1) as well as in just the engineering disciplines (1). Games having less than 1 million subscribers are shown.

whereas the study by R. M. Ryan *et al.* (25) found that it was critical to enjoyment and motivation as well. Goals in games can often be reached by multiple routes [Gee's "multiple routes principle" (16)]. But, in these branching decision structures, the learner must navigate between choices based on a considered estimation of relative consequences. In other words, the learner must operate at the highest level in

B. S. Bloom's *Taxonomy of Educational Objectives* (26), "Evaluation."

The active, participatory style of learning in games also departs from the traditionally passive lecture [Gee's "active, critical learning principle" (16)]. Game-based tasks often require the formation of hypotheses, experimentation, and discovering the consequences of actions taken; in other words, they are very similar to the inquiry-based learning lauded by science educators (27). Increasingly, game activities are multiplayer in design, meaning problems are set up to be solved in teams. Anywhere from a handful up to 40 players interact at a time via text or voice, sharing strategies in the pursuit of game goals and learning from each other as they engage in the activity. In this context, the teacher becomes a "wise guide" who participates alongside the students. Although no game-based data are available, classroom studies show that collaborative learning yields, on average, a 50% improvement over solo learning (28).

Finally, with all else being equal, games invite more time on task. Teenagers commonly spend 5 to 8 hours per week playing games, and this equals or surpasses the time spent on homework each week (29). B. D. Coller's racing car game, designed to teach numerical methods, resulted in twice the time spent by students on homework as a traditional class, with greater depth of understanding of the relations between concepts, and an overwhelming demand for the follow-up course (10, 11).

In contrast to the pedagogical and motivational elements found in games, some studies suggest that the lecture format is severely wanting. E. Seymour and N. Hewitt (30) chronicle near-universal antipathy to the undergraduate lecture experience, showing that 98% who leave science and engineering majors cite "poor teaching by faculty" as a major concern and that even 86% of those who stay say the same. R. R. Hake's metastudy (31) of 6542 students in 62 introductory physics classes demonstrated only a 17% SD in learning outcomes across lecture-based classes. In contrast, the same study showed that switching to any interactive mode of instruction (e.g., group projects, Socratic lectures, participatory demonstrations) easily improved learning outcomes in
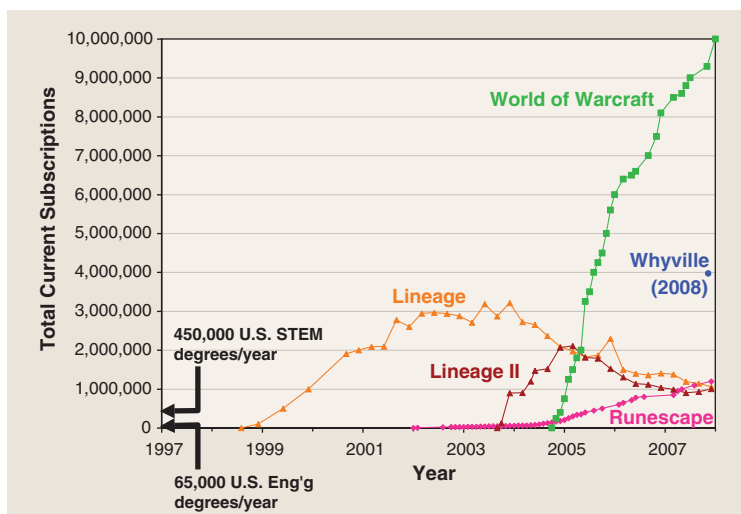
introductory physics by 108%. One could certainly argue that games are about the most interactive type of content that exists today. If video games are valid pedagogical delivery vehicles and they reach many more people than lectures, why do we not see video games adopted as the learning vehicle of choice? Cultural adoption lag exists, but we also face challenges of quantity, quality, and sustainability.

Quantity

It is often assumed that games with academic content are inherently uninteresting. Yet, 4 million children voluntarily play math-and-science–based exploration games on Whyville.net (*7*). In my opinion, most academically developed games suffer from infrastructural challenges rather than content challenges, with respect to mass adoption.

Examples include the lack of any distribution mechanism for the product, the lack of product discoverability, the prohibitive expense of content creation, the dearth of meaningful assessment (and therefore of consumer confidence in the product), and the lack of sustainable business models.

The first infrastructural challenge is the lack of any mechanism for distribution, sales, or marketing. Grants will not pay for these essential business functions that are required to reach audiences in the millions. Instead, academic games are often relegated to the office shelf or personal Web site of their creator as soon as the grant is over. One way around this dilemma is for a third entity—for example, a not-for-profit organization—to take on the business activities in exchange for intellectual property rights from the content creator.

Regarding the challenge of discoverability, academic game producers often use the Web as their distribution mechanism. However, three-dimensional (3D) content is not discoverable by search engines, which read text and text-based tags. For someone interested in capacitors, for example, Google cannot discover a virtual 3D capacitor in the middle of a game about electronics. Therefore, a key need in the area of 3D immersive games is the institution of a standardized metadata tagging system that allows users to locate appropriate 3D content with the use of common search engines. For the visually impaired who "see" 3D content only via voiced expression of tags, this tagging system is crucial. At present, there are multiple inconsistent tagging systems in use by specialized communities, but most games embed none of these.

Expense is also an important factor. User-created 2D content floods the Web. We can imagine a future in which the same is true of 3D content, and this richness of content could spur a concurrent, expanding user base of 3D games, large and small. However, the reason that 2D content is so cheap and easy to generate is the fact that almost all of it can be easily repurposed: copied, pasted, and moved from one application, document, clip-art bank, or Web site to another. In contrast, 3D content has no standard file format and thus has a limited ability to repurpose content between applications. Moving to a common file format for 3D objects—Collada and/or X3D (*32, 33*)—would greatly reduce graphics development costs, moving high-quality video game creation into the academic/home-user price range.

Quality

The ability to distinguish between a high- and low-quality product will be essential to the growth and credibility of game-based learning as a field. However, the first step in delivering quality is to be able to measure it. Assessment data are notoriously expensive to obtain, typically costing as much to develop as the original game. Few funders are willing to bear this double cost. To address this issue, the Ewing Marion Kauffman Foundation (*34*) has begun investigating the possibility of creating a software infrastructure to automate certain assessment tasks, thereby standardizing assessment across different games, lowering the cost of assessment per game, and making it more likely that researchers and funders will engage in assessment activities. Automated assessment is surprisingly advanced in certain areas: For example, automated essay grading is now nearly identical to human essay grading (*35, 36*).

Games may also extend assessment into new areas. Whereas we say that we value 21st-century skills such as problem solving, teamwork, communication, and leadership, these essential traits are nowhere to be found on a modern transcript. An attractive dimension of game-based assessment is the potential to track sequences of user actions and communications, then map these onto higher-order

Fig. 3. Protein-sized drone flying over macrophage surface in *Immune Attack* (*9*). The player is required to call neutrophils by using the drone's ray gun to activate CXCL8 release.

Table 1. Learning outcomes of several games compared to lecture on same material.

Game	Topic	Audience	N (study size)	Learning outcome over lecture	Reference
Dimenxian/ Evolver	Algebra	High school	193	7.2%	(*37–39*)
Geography Explorer	Geography	College	273	15 to 40%	(*40*)
NIU Torcs	Numerical methods	College	86	2× more time spent on homework, much more detailed concept maps	(*10–11*)
River City	Ecology/ biology	Middle/high school	≈2000	15 to 18%, on average	(*13*)
Supercharged!	Electrostatics	Middle school	90	+8%	(*41*)
Virtual Cell	Cell biology	College	238	40%, on average	(*40*)

skills and abilities. For example, in the case of problem solving, one can easily measure how often a user attempts a given problem. Attempt frequency (especially if each attempt is different) correlates highly to improved problem solving. Similarly, by monitoring users' keystrokes while they navigate search engine results, we can distinguish between hypothesis-driven searches and random searches, another key indicator of advanced problem-solving skills.

Sustainability

The last major hurdle in expanding the use of game-based learning is arriving at sustainable business models. Academic game development, which depends on living from one grant to the next, is inherently unsustainable. However, if funders could lay the foundations in an initial grant, the same learning materials could transition to profit-generating models that could be used to expand the material's reach after small-scale academic development is completed. These models could include corporate sponsorship, dual pay (free to some, but a fee for others) or sliding-scale fee models, subscriptions, site licensing, and the sale of virtual goods (e.g., virtual clothing to be worn by the player's in-game character, downloadable wallpapers, electronic books that give game hints). Other business models could include leader sales to countries with nationalized education systems and hence centralized buying power, partnerships with commercial game distributors, and microcredits for microknowledge (a far-future economic concept wherein a user would pay, say, $0.99 to learn the Pythagorean theorem via a small educational module, in exchange for a math mini-credit that could aggregate with other mini-credits toward a degree). To my knowledge, none of these methods has yet been used to sustainably support academically developed games, with the possible exception of corporate sponsorship, which has supported the growth of academically developed but for-profit–operated Whyville.

Summary

Although the field is still in its embryonic stages, game-based learning has the potential to deliver science and math education to millions of users simultaneously. Unlike other mass-media experiments in education (e.g., TV, Webinars), games are a highly interactive medium with many key attributes shared with sophisticated pedagogical approaches. Large-scale adoption, however, still awaits key infrastructural developments to improve quantity (of users), quality (of product), and sustainability (of business models).

References and Notes

1. *NSF Science and Engineering Indicators 2008*, NSB 08=01, 08-01A (National Science Board, Arlington, VA, 2008), appendix table 2-27.
2. www.worldofwarcraft.com (United Nations World Food Programme, Rome, Italy).
3. B. S. Woodcock, MMOGCHART.COM **23.0** (2008), available at www.mmogchart.com.
4. www.food-force.com (United Nations World Food Programme, Rome, Italy, 2005).
5. United Nations World Food Programme, various press releases, available at www.food-force.com/index.php/ press/releases/undefined.
6. www.whyville.net (Numedeon, Pasadena, CA).
7. J. Bower, paper presented at the Grantmakers for Education Annual Meeting, Baltimore, MD, 20 to 22 October 2008.
8. Subsequent personal communication from J. Bower detailed the exact percentage of female Whyville users, at 68%.
9. http://fas.org/immuneattack/ (Federation of American Scientists, Washington, DC, 2007).
10. B. D. Coller, paper presented at the Serious Games Summit, Washington, DC, 30 to 31 October 2006; available at www.ceet.niu.edu/faculty/coller/research.htm.
11. Also, for more detailed information on concept map results, see B. D. Coller's related work at www.ceet.niu. edu/faculty/coller/images/handsOnPreprint.pdf.
12. http://muve.gse.harvard.edu/rivercityproject/ (Harvard Univ., Cambridge, MA, 2007).
13. D. J. Ketelhut, C. Dede, J. Clarke, B. Nelson, paper presented at the 2006 AERA Annual Meeting, San Francisco, CA, 7 to 11 April 2006; available at http://muve.gse.harvard.edu/ rivercityproject/documents/rivercitysymping1.pdf.
14. R. T. Hays, "The Effectiveness of Instructional Games: A Literature Review and Discussion" (Technical Report 2005-004, Naval Air Warfare Center Training Systems Division, Orlando, FL, 2005).
15. J. J. Vogel *et al.*, *J. Educ. Comput. Res.* **34**, 229 (2006).
16. J. P. Gee, *What Video Games Have to Teach Us About Learning and Literacy* (Palgrave MacMillan, New York, 2003).
17. D. G. Oblinger, *J. Interact. Media Educ.* **2004**, 8 (2004); available at www-jime.open.ac.uk/2004/8/.
18. M. Begg, D. Dewhurst, H. McLeod, *Innovate: J. Online Educ.* **1**, issue 6 (2005); available at http:// innovateonline.info/?view=article&id=176.
19. J. Willinsky, *Science* **323**, 39 (2009).
20. A. C. Peck, D. Detweiler, *Hum. Factors* **42**, 379 (2000).
21. S. Barab, M. Thomas, T. Dodge, R. Darteaux, H. Tuzun, *Ed. Tech. Res. Dev.* **53**, 86 (2005).
22. S. Barab *et al.*, *Sci. Educ.* **91**, 750 (2007).
23. A. C. Graesser, N. K. Person, *Am. Educ. Res. J.* **31**, 104 (1994).
24. D. Druckman, R. A. Bork, Eds., *Learning, Remembering, Believing: Enhancing Human Performance* (National Academy, Washington, DC, 1994).
25. R. M. Ryan, C. S. Rigby, A. Przybylski, *Motiv. Emot.* **30**, 344 (2006).
26. B. S. Bloom, Ed., *Taxonomy of Educational Objectives* (Susan Fauer Company, Chicago, 1956).
27. J. D. Bransford, A. L. Brown, R. R. Cocking, Eds., *How People Learn: Brain, Mind, Experience, and School* (National Academy, Washington, DC, 2000).
28. D. W. Johnson, G. Maruyuma, R. Johnson, D. Nelson, L. Skon, *Psychol. Bull.* **89**, 47 (1981).
29. L. J. Sax, J. A. Lindholm, A. W. Astin, W. S. Korn, K. M. Mahoney, *The American Freshman: National Norms for Fall 2001* (Higher Education Research Institute, Los Angeles, CA, 2001).
30. E. Seymour, N. Hewett, *Talking About Leaving* (Westview, Boulder, CO, 1997).
31. R. R. Hake, *Am. J. Phys.* **66**, 64 (1998).
32. www.collada.org (Khronos Group, Beaverton, OR, 2006).
33. www.web3d.org/x3d/ (Web 3D Consortium, Menlo Park, CA).
34. www.kauffman.org (Ewing Marion Kauffman Foundation, Kansas City, MO, 2008).
35. D. Hutchison, *Br. J. Ed. Tech.* **38**, 977 (2007).
36. D. E. Powers, J. Burstein, M. Chodorow, M. E. Fowles, K. Kukich, *Comparing the Validity of Automated and Human Essay Scoring* (GRE No. 98-08a, ETS RR-00-10) (Educational Testing Service, Princeton, NJ, 2000).
37. M. Kebritchi, thesis, University of Central Florida, Orlando, FL (2008).
38. M. Kebritchi, A. Hirumi, H. Bai, "The Effects of Modern Math Computer Games on Learners' Math Achievement and Math Course Motivation in a Public High School Setting," available at www.dimensionm.com/docs/ UCFResearch_Brief_June_202008.pdf.
39. Reference (*37*) is the basis for (*38*).
40. P. McClean, B. Saini-Eidukat, D. Schwert, B. Slator, A. White, in *Selected Papers from the 12th International Conference on College Teaching and Learning*, J. A. Chambers, Ed. (Center for the Advancement of Teaching and Learning, Jacksonville, FL, 2001), pp. 111–118.
41. K. Squire, M. Barnett, J. M. Grant, T. Higginbotham, in *Proceedings of the 6th International Conference on Learning Sciences*, Y. B. Kafai, W. A. Sandoval, N. Enyedy, A. S. Nixon, F. Herrera, Eds. (UCLA Press, Los Angeles, CA, 2004), pp. 513–520.

10.1126/science.1166900

Laptop Programs for Students

Andrew A. Zucker[1] and Daniel Light[2]

With the continuing decline in costs of technology, programs are proliferating worldwide to put networked laptop computers into the hands of millions of students on a routine basis. The reasons policy-makers support these programs are based on economic arguments, equity concerns, and widespread interest in education reform. Studies of laptop programs in schools report that they increase students' engagement in school, improve technology skills, and have positive effects on students' writing. However, evidence of the effectiveness of large-scale laptop programs in other learning domains is scarce. Research in many nations suggests that laptop programs will be most successful as part of balanced, comprehensive initiatives that address changes in education goals, curricula, teacher training, and assessment.

Interest in providing laptops to schoolchildren has been growing for more than a decade, with a school in Australia beginning what may have been the first such program in 1990 (*1*). Traditional manufacturers now offer many laptop models costing under US$900. In addition, less-expensive laptops especially designed for children and schools have become available, including the XO computer designed and distributed by One Laptop Per Child [a spinoff of the Massachusetts Institute of Technology (MIT) Media Lab] and the Intel Classmate personal computer (PC). Ultra-