THE DECLINE OF THE BEST[1]

Claremont Graduate University
& Western Michigan University

Introduction

This is a study of some of the ways in which technology relates to educational evaluation, with an emphasis on interactions with globalization. Its particular and closely linked foci are: (i) an evaluation of *technology education* (an exercise in curriculum evaluation); (ii) an evaluation of *educational technology* (a case study in technology assessment); (iii) an evaluation of the present and future *use of technology in evaluation* (an aspect of disciplinary evaluation). In each of these efforts we begin with a look at the historical or pre-historical origins of the entities under consideration, and consider the US situation in some detail before attempting some global reflections; and in each case we look for some lessons learned to guide future development.

It's a commonplace that the evolution and globalization of technology has made it possible for our species to destroy all life on earth, and it's equally clear that the biological evolution of our species has not yet stabilized the avoidance of that result as a dominant characteristic in the population, although it is now a key survival characteristic. The discussion here is intended to display some evidence for a critical intermediate conclusion, namely that evaluation, itself now globally seeded through some technological developments that are discussed here, can provide an essential tool for identifying solutions and pseudo-solutions to the problem of human survival, although this tool has not so far been optimally employed.

This conclusion is partly based on a case study that shows how the United States has failed to use expert evaluation to save its own educational system from serious decline—and the prospect of continuing decline—with respect to its coverage of technology, and its use of technology, a failure that inevitably dooms its technological edge, once the best in the world. In this failure it is not very different from most of the other leading industrial nations, who face a similar crisis, although perhaps not so immediately. Unfortunately, from the global point of view, this failure has rendered us unable to pass on the lessons we should have learned earlier than anyone else, lessons that would have helped the rest of the world progress at the rate and on the scale needed to provide substantial security against disaster.

---

[1] One section of this, on technology in evaluation, is a chapter in the forthcoming ***Handbook of Educational Evaluation*** (Sage, 2009). All of it has benefited from comments from the editors and reviewer for that book, for which this turned out to be much too long.

1. Meanings and Origins

In order to do an evaluation of anything—for example, technology—we frequently need to begin by clarifying the concept itself. In this case, that undertaking, often a mere formality, turns out to constitute a major problem. But it isn't a problem to be solved *before* we begin the evaluation; it's the first problem *in* the evaluation. Evaluation has a key role to discharge in the clarification of concepts whose application is under examination, something which is obvious enough in the case of concepts like 'just war' or 'justifiable torture' but no less essential with 'science education,' and 'technology.'

Let's begin by thinking about the earliest stages in the history of technology in the hominims.[2] That topic in texts, and in search engines, usually brings up something about flint tools or stone grinding basins, but to think of them as the earliest technology is a long way from the truth. They are merely the best *still-existing evidence* about early technology: common sense (and comparative anthropology and primatology) makes it almost certain that there would have been bone and wood tools, and shaped gourds, in use long before there were any handmade (as opposed to found) flint or stone tools.[3] Consideration of these pushes the birth of technology back well beyond 2.5 million years, not just the two million since the earliest hominim, *homo habilis*, used stone tools.

But even the oversimplified 'stone tools' concept of the birth of technology is nearer the truth than the one we find in the Oxford dictionaries—and many other references—that define technology as a branch of applied science.[4] The term "technology," unless you accept its recent corruption, has traditionally always been taken to refer to functional artifacts[5] and usually also to the body of skills involved in making and/or using them. Since technology, as any glance at a history of technology will reveal, is *always* taken to have been around since at least the stone age, and science is *always* taken to be just a few thousand years old, the idea that technology is applied science is absurd—it involves an error of well over a million years.

---

[2] That is, the sub-group of the hominids now seen as clearly ancestral to us, which excludes the australopithecines. (The best up-to-date general coverage of this and many other topics in this essay is in Wikipedia or Google, and only when more detailed information than they provide is required is a more specific reference provided here. (For almost every topic here, both are greatly superior to the Britannica, which was compared on each concept).)

[3] There are also some less tangible artifacts, of which domesticated fire is one of the most important, and one whose distinctive footprint is hard to identify; however, it probably did not originate before 1.4 million years ago.

[4] E.g., "the application of scientific knowledge for practical purposes" in the Oxford American Dictionary. The error of excessive restriction in this definition is roughly equivalent to defining history as the record of men's accomplishments, i.e., forget women, Krakatoa, and the Black Death.

[5] The term 'functional' is used to make the usual distinction of technology from art. Cave paintings are not normally described as technology because it's not clear they were created for instrumental purposes—the motivation was probably religious or aesthetic. Of course, later artists often *use* instruments like brushes or paint that are artifacts, but the *art itself*, as art, is usually covered under the history of art, not the history of technology. But this is a minor issue, and we eventually relax the requirement.

The error in the Oxford definition is extremely significant. It reflects one of the most impressive confidence tricks in the history of thought: the effort by Big Science, in post-World War II US, to convince government and the public that technology is applied science. This hugely successful slice of self-serving mythology, opportunistically based on the atomic bomb 'success case,' has meant that government educational funding provided with the intention of safeguarding the future of technology has mainly been diverted into science education, with some slight gestures towards technician training, neither of which covers the main content of basic technology education for the citizen, which should surely provide some basic understanding of the nature of the great enterprise on which we are completely dependent. Since science teachers themselves have rarely had any instruction in the nature or history of technology, which is hugely different from that of science—and often resent the idea that they should spend some of their limited curricular time on talking about it—this has meant some generations have been brought up in almost total ignorance about technology's method and requisite skills. Note that the training of *technicians* has continued, which is not the same as teaching about technology. There is also considerable *talking about* automobiles, the space program, and other technological achievements, which are frequently misrepresented as scientific achievements, much to the annoyance of many leaders in engineering education.

So our first step in any attempt at an understanding of the interaction of technology with other aspects of civilization comes well before we start looking at its relation to globalization and education, and uncovers a very important flaw in our society's understanding of technology itself. This is not a good basis for a society whose past and present success is built on technology, and is a disastrous basis for education about technology.

At this point we should, before continuing, consider a common response to this strong emphasis on a distinction between science and technology: the response that while technology was not, at first, applied science it is *now* properly considered, and defined, to be so. The problem with this approach is that technology hasn't changed its essence over the millennia; and that essence is totally different from the essence of science. Technology has always been the enterprise whose content and product is *artifacts* and whose method is experimental or industrial *physical construction*—by hands, voice, or electro-mechanical means. Science, on the other hand, is now, and always has been, a *branch of knowledge* whose product is systematic explicit knowledge beyond that included in common sense, made up of entities such as descriptions,[6] classifications, explanations, predictions, and generalizations, and whose essential method is *critical and analytical thought*.[7] Explanations and theories, etc., are all things that require and indeed are part of language, and all of which are entirely abstract entities, not physical products at all. So these two creative enterprises are not just different, they are *essentially* different; and the skills for creating them, indeed the interest in creating them, is entirely different. True, science—now that it exists—can

[6] Most recently, it has become clear that the descriptions may be evaluative as well as purely empirical, since it is clear that many of the most important results of scientific investigations are determinations of the merit (or lack of merit) of e.g., medicines, radiations, and treatment regimens—and also theories, lines of research, and classifications of all kinds.

[7] I do not list experimentation, often mentioned in defining science, since it was obviously absent from the early work in astronomy, as it is in most geology, botany, etc.

often suggest ideas that can be turned into technology, but the reverse is also the case. Applying science is not the essential nature of technology, even now, and many great technologists,[8] from Thomas Edison to the Wright brothers, to Steve Jobs, were not applying science when they made their notable contributions to technology.

In addition, it's a mistake to think of technology as normally the result of a single inventor's genius. It's usually a development from—a merging of—many ideas from many people. Nobody invented the knife you use at your dining table, or the table itself, or the house you live in, or the automobile you drive—although some notable inventions from Gottlieb Daimler and Henry Ford and a hundred others *contributed significantly* in the latter case. Even technologies that we were told in school were the results of one person's invention—Stephenson and the steam engine, for example—turn out to be far from that, once serious historians start looking into the matter. Stephenson made a contribution, but what he produced was commercially almost useless; many years of further development in the tin mines of Cornwall by unknown tinkerers were required before the engine of the industrial revolution existed.

The push to identify an 'inventor' of a technology comes from trying to apply the 'great scientist' paradigm, where the original discovery often really is just the idea when first conceived, uttered, or published. But technologies normally only said to *exist* after a long process that requires something to be developed, patented, manufactured, distributed, and used. There are indeed some exceptions—there are even cases of Nobel Prizes in science being awarded to individuals for inventing a technology (e.g., the bubble chamber)—but they are rare exceptions rather than the rule.

One last point on the differences. Sometimes, in this kind of discussion with scientists or philosophers, one of them will become exasperated and say, "Surely science and technology are essentially the same enterprise—they are both just critical problem-solving, both use the methods of hypothesis testing, evidence checking, and so on." Ah, yes, and so does solving the crossword puzzle in the Sunday paper, and cooking, and mathematics, and running a business. They all take smarts, and smarts has some vaguely defined common elements. But this is such a general resemblance that it obliterates most useful distinctions. So one might reply, "At that level, there's no difference between a physicist and a biologist." If the answer is, "Of course there is, but it's a subject matter difference" then one should reply; "Then that's the distinction between science and technology; and it means being good at one doesn't make you any good at all at the other. They are different subject matters, basically artifacts vs. explanations, and a more detailed look at their methods shows big differences there, too. What we're doing here is spelling out those differences, and if there's a reason why in science education we try to cover both physics and biology, there's at least that much reason for having equal time for science and technology."

It's time to return to our look at the early history of technology, because getting clear about this uncover some other surprises. Our species is currently dominant—even if only for a

[8] The term "technologist" is not sharply defined in English; it is used here to mean someone who makes or is working competently at making a significant contribution to technology. This may be an engineer, or an inventor, or an illiterate tinkerer; but such a person is quite different from (not superior to) a technician, in skill-set, goals, and activities.

fraction of one percent of its history—but it achieved that status with a genetic endowment that is a mere 1.3% different from the chimpanzee. The dominance appears to come from one biological advantage, emergent from and nurtured by a particular series of social interactions, and resulting in one great artifact. That advantage was our slightly greater capacity for prelinguistic problem-solving, which eventually and very slowly led to the key artifact of a structured language (by contrast with a set of signals), which in turn—no doubt with some assistance from the parallel explosive tripling in cranial capacity[9]—made possible both symbolic reasoning and the creation of an autonomous body of (pre-scientific) knowledge. This 'public body of knowledge' (sometimes called 'intersubjective knowledge') about our environment, and about and tailored to our own cognitive and psychomotor abilities, could be passed along and accessed via speech or writing without access to a person in whom it was tacit. This facilitated the 'globalization of knowledge' and eventually made possible the present widespread distribution of considerable power to manipulate that environment, even though that power is still very limited with respect to natural disasters (and some we create ourselves).

Despite the enormous boost that this reservoir of linguistically encoded public knowledge gave to human learning, it is suggested here that the implicit component of our knowledge and skills, often called tacit knowledge/skills, which is fully teachable and learnable, may still be almost as important as the explicit one. Keep in mind that during most of the evolution of homo sapiens, including the long prelinguistic and early language epoch, tacit knowledge and tacit reasoning were the only or the main source of knowledge, and good enough to out-survive almost anything else around—and to generate much technology including language refinements. Tacit knowledge and skills, passed on by parental guidance, not by talking, are what we often see in mammalian problem-solving behavior and of course, we are still genetically almost the same organisms that we were in that epoch…[10]

Technology is a social or cultural phenomenon (as is globalization); we don't speak of a technology as really existing until it has come into use in a community. So we only use the term hereafter to mean the set of functional (and perhaps also decorative) artifacts *in use by a community at a given time*, and the knowledge, skills, and equipment to produce and use them.

Now proto-globalization—the dissemination of culture—occurs in its basic form without the emergence of language and probably without the help of any other substantial technol-

---

[9] Actually, rather more than tripling, but with some increase in body mass as well; nevertheless, the ratio of brain size to body size was almost tripled. The exact nature of the causal relations here is still much debated, but it's arguable that a feedback loop operated between the advantages from language and brain size. That is, when the brain size increased enough to support more sophisticated language structure, the survival benefits from the better language increased enough to make it more likely that that variation remained in the gene pool.

[10] These dots indicate the first of several places where there has been an excision of about a thousand words or more, at some cost to completeness and/or continuity. The penultimate draft of this paper ran 25,000 words, and even with a great manifestation of charity by the editors, for which I am exceedingly grateful, 8,000 words had to come out. The uncut version is online as "Decline of the Best/Draft" (http://homepages.wmich.edu/~mscriven).

ogy.[11] Education, and at least this primitive version of globalization, are thus perhaps the oldest of the four concepts we are connecting here, though evaluation is arguably as old as education, since teaching is largely, although not always, a matter of teaching how to do something *better*, which involves *learning to distinguish better from worse*, i.e., evaluative skill (implicit at first).

At a certain stage in global history, the pace picks up again, and this time technology provides a further series of mighty accelerants. It is these contributions of technology that eventually distance homo sapiens from the pack. Of course, the primates, and especially the bonobos, select or make and use tools (including stone tools), without a real language, and may perhaps—this is not entirely obvious—use them more intelligently than the beavers, spectacled bears, and weaver finches use their building materials. We can reasonably suppose our ancestors did this too. And many creatures educate their young and their peers without language. But, as already stressed, our ancestors developed just one tool that gave them an edge—perhaps even leading to a larger brain—on which depends almost every other advantage their descendants have: they created a general purpose spoken language.

Although language is not a substantial material object in its original and essential spoken form (but then nor is fire), it is an artifact, and hence its development should be regarded as a part of technology. Even historians of technology have often not looked carefully at massless things created incrementally, across epochs, by communities rather than by individuals. Yet it is from that community of historians that we have come to realize the absurdity of crediting most technologies to individuals.

2. Proto-Globalization

The emergence of language was not exactly a Eureka event, since its creation may have spanned a hundred millennia—not just a hundred centuries, but a hundred millennia. But it was the Great Accelerant of progress, as conventionally defined, because it meant that knowledge could be communicated without demonstrations by the cognoscenti, and this opened up to dissemination vast ranges of knowledge about the world which cannot be demonstrated because they are not skills but simply information.[12] This kind of knowledge, and not just skills, could now be carried by the trader, the traveler, the seaman—and additions to it could be brought back by them when and if they returned home. So now a richer kind of globalization[13] became possible, the globalization of our explicit knowledge—or to

---

[11] It is useful to distinguish active and passive technology; an active technology is one where the design skills are also still present, passive technology refers to the relics of earlier, possibly still functional, technologies that are no longer within living powers. For example, the pyramids were part of Egypt's passive technology at the beginning of the Christian era, but Egypt no longer had the design or building skills that created them.

[12] As well as knowledge that is *about* skills, that is, ones the knower can describe although not personally demonstrate, such as the use of an atlatl (spearthrower) or knapping (skilled flint-tool making). This opens the possibility though not the virtual certainty of emulation, and thus language spread both knowledge and skills.

[13] This term has no precise generally accepted single meaning (as a 'define' command in Google makes clear) but we here take the core meaning to be represented in the following definition: orga-

be more precise, our beliefs—about the world. Correspondingly, education's physical reach now expanded a thousandfold—and in moving those boundaries outward, the business of discovery (i.e., the addition of new-to-anyone knowledge) expanded with the boundaries, paid for by trade and motivated at first by the hope of wealth rather than the quest for knowledge.

The next great accelerant was the separate technological development of the *written* language, since it meant that knowledge could be passed on without having to pass through the head and voice of an intermediary. Hence even specialized knowledge could be globalized, knowledge that could not be carried in the head of the differently specialized human trader and traveler.

Of course, written language was not the first kind of informative inscribing that traveled well; drawing and painting probably preceded it, and were presumably used to illustrate spoken accounts of skills as well as wonders seen.

Now discovery and communicating new facts about the world is a precursor of science, although it is a long way from the systematic knowledge that science involves, let alone any experimental or theoretical discipline. So the explorers led the expansion of a kind of pre-scientific knowledge, and millennia later an emerging world of science waited for their reports with great interest. But as technology itself developed, the need for the raw materials for manufacture as well as for ornament became increasingly important, so the miners followed the explorers, and the chemists emerged to refine the miners' searches. Much of this required or was greatly facilitated by the written as well as the spoken language, though the earliest miners—of flint, long before the metals—whose lengthy trade routes in Europe we have now traced, may have operated without either.

In 'globalization by emigration' that began to spread culture in parallel with conquest, hunting, and trade, the mode of transport was at first the feet, perhaps next the beast of burden. The use of the latter required the technology of harness and packs, and the travois, although not wheeled vehicles until much later. But along the lakes, rivers, and coasts, travel was sometimes faster as rafts, coracles, and canoes emerged and eventually became oared boats and finally sailboats; this was marine technology in the service of globalization. Travel inland, often over mountain passes, and eventually into the northern latitudes, required substantial development of both clothing and habitation. So, possibly before spoken language, technology was serving early globalization, which we'll call 'slow globalization.'

3. Technology in Education: (I) Curriculum

This topic concerns part of what is here called 'technology education' by contrast with 'technical training.' There is extensive technical and engineering education in the US, to maintain and to some extent develop the machinery of civilization, and much emphasis on science education, but there is little attention to two other matters: (i) in general education,

---

nized (and usually centrally controlled) social influence—especially cultural, commercial, and/or political influence—across national boundaries, on a very extensive if not literally global scale, often at a very high level of impact. For example, the commercial strand of the influence of globalization alone may overwhelm even strong political and ethical considerations.

which has some coverage of science, little time is spent on the nature of the very different beast on whose back every citizen rides; and (ii) no career track is there for the technologist to follow. Instead, the student is offered only science education, on the one hand, and technician education on the other. The missing elements cost us dearly. On the one hand, we lost our leadership in the automobile field, where Japan was training technologists to redesign cars for not needing maintenance; on the other, we raised generations with no understanding of the tiger they were riding on.

Signs of this ignorance are everywhere and cumulatively a matter for deep concern. As a telltale example, in itself of no great importance, discussion of 'technology' in educational circles, and in the flyers from publishers, is today always just a shorthand reference to computers, clear evidence of a failure to realize that computers are a mere current event in the sweep of educational technology. One might as well use the term 'business' to refer only to banking, or 'war' to refer only to wars between navies, or 'education' to refer only to mathematics instruction. Education is totally dependent on a hundred other technologies, half of them continually and radically evolving, and has been thus dependent since before there was a schoolhouse and a slate, a seat, or a sheet of paper. We just take those for granted as part of history or the background, never having had them called to our attention as matters for study and concern, for management and costing, for research and development—in short, as respectable subjects for careers, consideration, and care.

A more serious phenomenon is the example set in education by the National Science Foundation, chartered as the official guardian of technology as well as science but fundamentally confused about the difference. It has major programs that it calls technology education programs when they are simply technician training programs. Technicians are mechanics operating machines and the other apparatus of technology or science—essential players, but neither scientists nor technologists. But the term 'technologist' as used here parallels and is essential to match our use of the terms 'scientist' or 'mathematician:' Technicians with talent sometimes become technologists, yet that transition is not automatic, any more than they are automatically scientists. The main problem is that the programs in question involve no serious *education* about technology or how to create, fund, relate to, or improve it; they are just *training* programs, i.e., skill-building efforts in the use of particular technologies. This is not the way to develop technicians into technologists. We need both these specialists: and to develop the latter, who are absolutely vital for the country, we need serious education in technology. Yet our leading agency in charge of the domain isn't even clear that technology education is not the same as technician skill building.

Another telltale of the fundamental conceptual failure here is that there is no curriculum subject for technology in the US school system, although there is an excellent national curriculum component on that topic in, for example, the UK.[14] This lack in the US is of course due to the national illusion that technology is applied science, reflected, these days, in many of the dictionaries and other reference works.

Good historians have argued that science is essentially a byproduct of technology; none of them have ever believed the reverse. Today of course, many interesting and newsworthy events in technology spin off from science, but that's partly a sign of what journalists make

---

[14] This is their General Certificate of Secondary Education subject, Design & Technology.

of them and partly a sign of overlap in this latest phase. But journalists must all know—they have made good stories out of it in the past—that the microcomputer, and its breakthrough software the spreadsheet, and the breakthrough application of Google, amongst dozens of other recent examples, were invented by students or others with little or no training in computer science. And it only requires navigating a single course or book on the history of technology to know that many of the great inventions in the history of technology were made by people with no training in science. Since the time of Archimedes, that mighty master in both fields, it's hard to find a score of great scientists who would qualify as great technologists or vice versa. After all, the two enterprises are totally different in aim and detail, as different as an idea from Einstein is different from an iPod.

If we were writing a textbook as part of technology education, let's consider for a moment just what would be in there about the creative process that establishes a breakthrough in technology compared to science. In science, it's easy enough to describe this—the idea just has to be written down or spoken in enough detail to be identifiable as a new and important contribution to knowledge. The classic example is the physicist Robert Hooke who wrote down and published the basic law of elasticity named after him in four Latin words—Ut tensio, sic vis (the deformation of an elastic body is directly proportional to the magnitude of the deforming force).[15] The best-known modern example is Einstein's law of the equivalence of mass and energy, $E = mc^2$.

But in technology, what the inventor famously scratches on the back of an envelope is a diagram of a thing, not a statement, and it's a long road from there to a thing that can be called a new technology. The road involves at least fifteen steps. (i) First, of course, a working model of the new thing in physical form has to be made, which may turn out to be very difficult or impossible. (This is where training in design, sometimes said to be the key element in training for technologists pays off.) Then there is the rest of the iceberg—the long and arduous development process, i.e., (ii) making enough models for serious testing in simulated long-term use, and doing that testing; (iii) redesign and improvement cycles coming out of that cycle of testing; (iv) blueprinting a version and having an industrial engineer calculate manufacturing costs; (v) getting an accountant's estimate of costs and probable net profit, if successfully marketed, and further capital needs; (vi) informal secret checking of past patents for originality, followed by (vii) the official patenting process; (viii) the search for production funding, once the cost of that is calculated; (ix) finding a manufacturer or premises for manufacture; (x) staffing and running, or supervising, the actual manufacturing process, which will almost always require (xi) further modifications and testing of them; (xii) a successful marketing effort of the appropriate size, sometimes necessary on an very large scale (where a market has to be created not just exploited); plus (xiii) setting up, staffing, and running a servicing and warranty management system; and (xiv) the staffing and management for oversight of all of the preceding (which cannot all be done by the small staff of an investment banker). Finally, (xv) only if the marketing is followed by suc-

[15] Actually, he first published it as a puzzle in 1676, as an anagram made up by jumbling the letters in the Latin sentence, only announcing the solution two years later, a rather odd procedure. However, since I was first told that story in physics class when I was ten years old and, despite my Latin being very poor, I have effortlessly remembered that sentence across seven decades, Hooke may have known something about pedagogy as well as elasticity.

cessful sales can we claim that the technology has been adopted, and only if it's adopted can it be called a new technology (because the definition requires use by a community).

Putting it another way: inventing a bright idea in technology is hard enough, but it's usually the easy part of advancing technology. Only if you understand the rest of the process will you be minimally prepared for understanding what it takes to make a successful contribution to technology, and only if you thoroughly understand it will you have a minimal preparation for a career as a technologist. Once you've read this, and if you are at all familiar with our present educational system, you will realize that we are not systematically producing technologists at the moment, only accidental and badly equipped ones. Worse, we are not producing citizens who understand what technologists need in order to make the contribution the country needs to maintain employment and keep up with the changing scene in the rest of the world. That means we are failing to prepare voters for the many current situations where they need to vote intelligently in campaigns where major issues concern technology e.g., our dependence on imported oil, and what it would take to eliminate or massively reduce that via alternative energy sources; our need to reduce and recycle waste and how to do that; how to reduce the death rate on the roads by better safety technology in cars, driver training, and road design; whether we need to 'keep up with the Jones' in atomic or conventional warfare technology; whether and on what scale to invest in more subsidies for new technology start-ups or technology education or technician training to handle unemployment; and a dozen more. When elections, all the way from presidential to school board, hinge on such issues and the education system hasn't even taught the difference between technology and science, let alone the skills and knowledge to follow the arguments about these issues, we need to rethink the curriculum.

The following is a real world example of many that could be given to illustrate why the citizenry needs to be knowledgeable about technology even if very few of them will ever become technologists. When private capital is not forthcoming to support economic development in a particular area, through lack of experience, governments understandably consider stimulating investment in that area. When I was working in Western Australia, a government was elected on a platform that included making that effort, because it was a chance to create very large numbers of jobs. 'WA,' as it's called, has been found to have more proven iron ore and diamonds than any other country in the world, but it has almost no processing industry, so it merely exports the raw materials. Setting up some successful 'downstream processing' industries would mean great economic growth, so the government proposed creating a small industrial park and providing some short term subsidies and tax breaks to get infant industries of this type started. A few of the first beneficiaries of this scheme failed, and the opposition party attacked savagely, along the usual lines of 'government should stay out of business,' 'waste of taxpayers' money' etc. Should the government persist or give up?

Clearly, the decision rests on a factual issue—what rate of success can one reasonably expect from those who are experts at doing this kind of investment? No-one in WA had the answer, or at least no-one was willing to provide it. But the answer, as near as one can get it, given that no-one likes to admit errors, is that in a country like the US where this kind of venture is often funded privately and very smart investors are involved, the best you can hope for is about one success in eight carefully chosen ventures of this kind. So a few failures is not a sign of bad investment decision-making, it's par for the course. Everyone should know that figure, or a more accurate replacement for it, so that a perfectly sensible effort to stimulate economic growth will not be abandoned for bad reasons.

But that kind of state policy issue is not the only type requiring better technology education. It's common for amateur inventors to mortgage their house, or let their parents mortgage *their* house, to fund what seems like a bright idea when external funding is not forthcoming. Knowing this failure rate would make people much better informed before taking such risks.

The message here is fairly clear: it's attractive to put it by saying "Technology is business, science is not business, it's just science." That's not quite right—science is often also business or partly business, and technology resembles science in minor or very general respects. But it's close to the truth, and closer than most people realize. People, including eminent scientists, coming up with an idea for a technological breakthrough are often shocked to discover that funders want 51% or more of the ownership in return for an investment. "But the idea is *all* mine," they exclaim, at meetings I have attended, "you're just doing the routine work." No: what the investors are doing is risking their money, a lot of money, at odds of 8 to 1 against, at best. If their judgment is not good, or if your idea is not really as good as you think it is, they lose their money. That means every winner they do pick has to pay for the investment of time and initial funding in seven losers as well as the winner—and, very importantly, for their time and expenses (expert consultants, travel, staff) in examining around 100 proposals before they find the 8 they are willing to back—before they make a penny, even if they're good at this. So they're not just greedy capitalists, they're sensible and expert capitalists, members of a group that frequently goes broke. Only if they were idiots would they let you hold 51% of the stock in a joint venture—which would mean you had full control over their money—given that you have no experience in running risky businesses, or in all probability, any other kind. You are just running on the 'science model' that makes the inventor of the new idea the owner of the benefits; that's not the way technology works or ever could work, since the idea isn't any use until it's a thing selling off the shelves in a market.

I hope that enough has been said to establish the real-world importance of improving technology education. In the first draft of this paper,[16] which was (inexcusably) twice as long, I went at some length into several other case studies which I'm going to skip here, in each of which I tried to show how some major business or policy failures in the US were dues to fundamental errors about the nature of technology. These included the collapse of Detroit's supremacy in the automobile world market; the charade of Homeland Security as a protection against terrorist attacks using airline facilities; the failure of the levees in New Orleans and on the Mississippi; the assumption that military hardware supremacy assures the success of an invasion.

We can conclude this section with a brief reference to a relevant experience by the author, reporting it as a participant observer. At the last stages of the national multi-year, multi-million dollar effort in the US to develop a 'K–12 science curriculum for the 21st century,' near the end of the 20th century (which was also intended to cover technology), I was called

[16] This is still available online; go to evaluation.wmich.edu, to the staff profiles, and click on mine for some online papers.

in by one of the participating agencies to review the final draft just a few weeks before public its scheduled public release, because they were a little nervous about its quality. I'll skip over the poor quality of the effort to identify scientific method and focus on the references to technology. In passing, I note that the content of the curriculum specifications for many grade levels in basic science were, by contrast, of excellent quality.

First, I note that of the one hundred experts on the multiple review committees that were part of the long development effort, only one was an engineer and none appeared to be successful inventors; so much for respect for technology in education. Second, I note here, as I pointed out in my report, that the definition of technology used had the rather remarkable but previously unnoticed properties of including Mother Theresa as a technologist and excluding the atom bomb as a technology. In other words, amateur night at the conceptual party. Hardly the best way to teach our maturing citizens about something on which their lives depend.

Now, let's reflect for a moment on what has been going on in this section. It has been an *evaluation* of the adequacy of the coverage of *technology* in public *education* in the US. To the considerable extent to which education in the US influences education in other countries, this bad example has been *globalized*. Fortunately for the exceptions, this globalization has not been complete, the UK not being alone in standing against our bad example. Of course, such exceptions make our mistake more serious for us in the global competition for leadership in technology. And the nation's political leaders—but not its business leaders—still seem largely unaware of the crisis. In spite of the overemphasis on science at the expense of technology, we are failing to produce either technologists or scientists at a level that meets our needs, and failing to support their work. Specifically, using NSF figures, we have cut funding for basic research in science and technology by 25% in the last 35 years; enrolment in science majors by our own citizens has followed almost exactly the same pattern; we are refusing green cards to allow overseas students graduating in these subjects in our colleges to stay on to work; and we won't allow overseas applicants to take the many unfilled high-tech jobs here. All of this has been repeatedly pointed out by high-level industry and educational study committees in the last few years.

It is surely time to consider whether the nation's evaporating leadership position in technology, and in technologist education, needs to be addressed more seriously at its basis in the curriculum vacuum about technology. This crisis is made more acute by its poor media coverage, a natural consequence of the absence of education of journalists about the subject, and exacerbated by the powerful public relations of Big Science, presenting itself as the solution.[17]

---

[17] These conclusions about technology education have some analogies in science education: there is no competent coverage in science education of education *about* science, or about its extent and boundaries. An attempt to cover scientific method is usually included, of poor technical adequacy; but there is even less attempt to cover such matters as the relationships between the sciences (the unity of science movement, reductionism/emergentism, etc.), the cost of scientific research, the validity of the method for evaluating scientific research (the lamentably inadequate peer review process), the great failures of science, the limits of science, the ethical issues inside science, or the relation of science to technology, engineering, mathematics, etc. There's nothing in that list that is beyond 10th graders' comprehension and interest, let alone juniors and seniors in high school. So

Another spin on the conclusions in this section would be to say that is shows that our society's shared abilities in technology assessment a.k.a. technology evaluation, are a long way short of competence. There's a deservedly famous effort on this in Stanford's engineering school but it's unusual, and variants needs to be available across whole campuses. In the next section, we'll address the converse topic—the impact of technology on evaluation—which is a rather more successful story. Then we'll turn to a combination of the two in a section on the failure of evaluation to deal well with educational technology.

A final twist in the role of this long defense of the importance of technology is that it may encourage evaluators to be more attentive to subdivisions of evaluation that program evaluators tend to ignore—product evaluation and technology education—although each is sometimes quite important in program evaluation, e.g., when the program produces materials, or uses expensive equipment, or might benefit from considering computerization of some of its functions; or when the evaluation of educational programs abuts on curriculum issues, or the evaluation of personnel abuts on their training.

4. Technology and Globalization in Evaluation

It's time to turn our direct attention to the relatively new profession and discipline of evaluation. Evaluation is a complex process, challenging for evaluators and anxiety-provoking for evaluees, and there are many very different accounts and varieties of it. But the aim of this paper is not to advance a particular account of it, it is simply to talk about how technology affects and might affect it. Even professional evaluation, some of which we hope was provided in the preceding section, is not a new practice, only a new profession. It is no more than *systematic critical appraisal using specialized tools and knowledge*—of facts, methods, and relevant standards—appropriate to the category of entity being evaluated.

"Critical appraisal," although a very loose description, does not cover all varieties of professional analysis such as statistical analysis or content analysis. It refers to analysis aimed at *determining the merit, worth, or significance* of the entity appraised (here called the evaluand). The need for specific professional training in evaluation has emerged as the need for a specialized level of training became clear in order to deal with an increasing number of difficult areas, such as large-scale program evaluation and quality control in manufacturing, and the increasing number of demands for detailed answers in those areas.

Evaluation as a profession began a thousand years ago, with the emergence of the Japanese proto-guild of sword-evaluators—an category of product evaluation—but only became self-conscious in the mid-20th century when program evaluators in the US formed the first association, a process since replicated in more than 50 other countries.

Of course, there were many people doing professional-level evaluation before then, even in the limited field of education, both in specialized areas like student testing and in many book reviews in educational journals. But evaluation of that kind was seen, correctly, as part of all professional approaches, not recognized—as was also true—as part of an extensive network of activities that had a substantial common logic, and much shareable practic-

---

it's not only technology that isn't covered, science is overprotected from the kind of critical review that is often said to be a key element in scientific method.

al wisdom. Rendering that logic and wisdom explicit has created the possibility of substantial improvement through self-critical study, the same process that had previously led to the emergence of science and technology as professional fields from anecdotal and artisanal activities. The development of evaluation from a profession into a discipline, as it is today, has in fact resulted from continued application of that self-appraisal and self-improvement process.

The questions of present interest that naturally arise at this point are: what effect has technology had on this newborn discipline, what effects might emerge in the future, and how does this relate to globalization? A list of suggestions follows, some explained here and some without detailed explanations when these can be found in appropriate references.

1. Technology has provided many *quantitative* analytic tools for the social sciences that are often of great importance in evaluation, especially software tools for computers used in statistical analysis (SPSS, SAS, etc.), and in financial analysis (Excel, etc.). These have become part of most professional evaluators' toolkits, and their existence is part of what justifies talking of a profession of evaluation.

2. Many analogous tools for so-called *qualitative* data analysis exist, although most of them are tools for converting certain kinds of data (often text or interview notes) into quantitative form, and dealing with it thereafter, and so might better be called hybrid or conversion software. They have a cadre of enthusiastic users, and developers, but their net benefits are less easily demonstrated than is the case with the quantitative software. This is because the skill required for coding is still very considerable, and the resultant data does not easily lead to firmly establishing significant conclusions not suggested by an equally skilled pass at the original material. Still, the results do sometimes turn up and confirm new hypotheses. A related example has come up in the multi-year Heifer International external evaluation project, where we have recently developed some customized tools in Excel that integrate qualitative and quantitative data, and that facilitate conversion of the combined data into graphical format (publication forthcoming in the ***Journal of MultiDisciplinary Evaluation*** (jmde.com)).

Much more is to be hoped for in this category, as the logic of qualitative data and pattern recognition software is further developed. For example, recent work on refining the 'modus operandi' approach to causal analysis suggests that this kind of approach might be programmable.[18] Various attempts are also being made to facilitate causal analysis by sophisticated statistical techniques that may move us further along that path.

3. Technology has provided ways to store project data in electronic databases that are easily and quickly searchable, and can be designed to facilitate the production of useful types of report including graphical reports. Some tools also exist that assist in suggesting novel implications of data, e.g., programs that automatically run randomized selection of variables for interaction testing. The most interesting future possibilities here are perhaps the 3D databases which exploit the human brain's high sensitivity to pattern recognition. Simple

---

[18] See "A Summative Evaluation of RCT Methodology & An Alternative Approach to Causal Research" Scriven, M., in ***Journal of MultiDisciplinary Evaluation*** 4.08 (at jmde.com)

examples display datapoints' clustering as spatial density in three dimensions selected from a dozen or more options. This approach has also been applied to concept mapping, and to mapping searchers' choices in Google information space.

4. Taking the last suggestions further in their own right, programming technology facilitates the use of graphical representation that can not only improve presentations (PowerPoint, etc.) but also facilitate understanding and the perception of significance and implications of data. While pseudo-3D graphics are common enough, and can be very illuminating (as well as sometimes confusing) it's easy enough to make true 3D, if that repays the trouble of distributing special glasses. We're still about 2 years from true 3D with lenticular screens (no glasses) for home video; it will appear in movie theatres late in 2008. Possibly more promising is the creation of videographics: showing 5 or 10 second videos to illustrate trends in a fourth dimension (time), which can be done with present technology. Although this has only speculative payoff for report presentation, its value is immense in another context, discussed below.

5. Photographic techniques for presentation of case study content and contexts, including video interviews, are important, as qualitative methodology has convincingly made the case for the importance of visual context in interpreting oral exchanges, through seeing both body language and environmental characteristics. There are more exotic evaluation tasks where this kind of equipment is invaluable, e.g., with infrared filters and motion sensors for detecting night visitors or guard movements.

6. Camera technology itself has taken several recent quantum jumps, each with some payoff for evaluators using or considering photographic media. (i) Several shirt pocket digital cameras in the low hundreds of dollars range were released in 2008 that take good short video clips with up to 5x tele lenses, excellent for portraits and also for making contextual vistas possible as backgrounds to audio records, which some of these cameras can also make. For example, the Canon SD 890IS (released 7/08) runs 6.5 oz, is the size of a pack of cigarettes (you remember them?), but with 10MP resolution and a 5x (optical, not digital) stabilized telephoto lens. (ii) The current small camcorders, not expensive, will fit in a jacket pocket, can do HD (high definition) and record in stereo. They won't give you high quality stills, but they make it possible to be unobtrusive and unassisted in getting motion records of service delivery or customer activity at a program site. (iii) If you want a camcorder for the *shirt* pocket, very simple to operate (e.g., fixed focus 1m to infinity), cheap ($100 or $150), low quality video (OK for emailing, looks poor on your TV screen), but remarkably good in poor light, go for the Flip Mino. For crisis/disaster assistance evaluation, a growing specialty with its new topical interest group (TIG) in AEA, or for our evaluations in remote villages from the Tibetan mountains to the arid forests of Peru, this will be invaluable for documenting claims of excellent or defective practices, shelters, animals, or supplies. (iv) In another performance category, the latest Casio camcorder (the Exilim EX-F1) enables you to take ultra-high speed video at a previously impossible price point (under $1,000), critical for some kinds of evaluation—e.g., evaluating athletic performance for judging or training purposes, or evaluating self-defense skill training, or for high-speed machinery/ordinance evaluation. (v) The Nikon D90, released in mid-August of 2008, at around $1000, is a rather bulky DSLR (digital single lens reflex) with a number of firsts for qualitative evaluators using photographic equipment: it's the first SLR with video, it allows 80 minute, HD videos with interchangeable lenses—something you can't get on a camcord-

er under $20,000. And it has auto focus and extreme low light capabilities (up to ISO6400) for still photography with professional quality detail in the shots. (vi) A week or two earlier, another breakthrough was announced by Microsoft, and it's free. This is something called Photosynth, and it allows anyone with any digital camera to do something that can be of great importance immediately in product and property evaluation, and of some importance immediately in personnel and some program evaluation; of great value in training evaluators when used for situation simulation and testing; and other applications will I think quickly suggest themselves. What it allows is 'spherical panoramas' with unlimited detail. For example, a real estate agent could stand in front of a property on the Kona coast in Hawaii and take about 300 pictures with any ordinary DSLR, using tele, normal, and macro lenses, and send them to Microsoft who will 'photosynth' them for you. Then you—the prospective buyer, in Manhattan—can project the result on a screen and zoom in to any level of detail on the garden, the façade of the house, the birds in the coconut palms overhead, the ants in the turf at your agent's feet, the surfboards on the distant beach, whales spouting miles out to sea. You want the closest thing to being there (with your binoculars), you can have it; in depth. Beats hell out of a postcard—especially in the Systine Chapel.

7. We should not fail to acknowledge the benefits for evaluation report writing of word processing programs, with the attendant spelling and grammar correction options plus instant dictionary/encyclopedia lookup, and near-instant translation of short passages in foreign languages, already in Word 08. The use of hypertext (cross-references built into docs so that one can link any point instantly to other locations relevant to the same point (definitions, encyclopedias, etc.), or to photos, or to fast-breaking news sites on the web); is still rare, but can be very useful especially in text and reference works. Voice annotation is available now, and voice recognition software, with the release in summer 2008 of Dragon Naturally Speaking version 10 is now at an extremely high level, far ahead of any of the three or four available competitors. It will now work quite well with accented voices, even without reading some 'training' passages as previously required, and it can handle editing commands easily, so that users with limited manual dexterity are able to dictate faster than anyone can type, and with very low error rates—one error a page is a common report from users after some experience, and that's well under the error rate of good typists. It must be noted that it won't work for recording interviews—it can only handle one speaker at a session because it heavily relies on an individualized database of their pronunciation.

8. A key tool for evaluators is the humble checklist. Its logic is not trivial (see the collection of checklists for evaluators in specific fields, and papers on checklists, at evaluation.wmich.edu), but its utility is almost universal across evaluation fields of application. The software for formulating and improving checklists, an extension of word processing software, is worth a moment's attention. This is an area of special interest for me, as the author of a book on word processing software and hardware, and the source of the inspiration—so he says—that led a student of mine, Ted Holm to invent hypertext.

It may seem odd to start by suggesting that something now seen as passé be reconsidered seriously, but I want to recommend it partly because it is useful far beyond the domain of checklists. I think we should not accept the commercial abandonment of what used to be called 'idea processors.' Some remnants of these are still available, as outliners, e.g., in Microsoft Word, and as standalone applications from many sources (search on 'reviews of outlining software' for up-to-date references), but the one most people know about, in

Word, lacks several features found in some of the full-blown idea processing software when it was popular. One of these is available in some of the independent products: graphical representation of the outline. This, especially the spoke and hub design, has some advantages over the linear design of standard outliners, because it gives sub-headings equal status (as spokes all bordering the same hub) instead of placing them in an order which suggests but does not ensure their relative importance. Another feature, perhaps dropped because no-one wanted to admit that it was useful, was the 'shuffling' option that, without losing the original arrangement, would randomly reallocate subs under headings, which quite often suggests new connections.

9. On the pure fact-hunting issue, there are now vast informational databases, and repositories containing multiple databases, instantly available that make access to comparative and baseline educational data either feasible or at least much faster and easier.[19] There are also general knowledge databases (Google, Yahoo, Wikipedia, etc.) as well as specialized educational databases (nces.ed.gov, etc.), and—still underused by evaluators—the huge image databases of which PicLens is by far the most spectacular and versatile. These are rapidly acquiring general question-answering capability (ask.com is slightly in the lead here) and we may expect better voice-driven interfaces shortly, although some human-serviced ones are currently available without charge, e.g., from Google. In either text or voice version, the big new advantage we are looking to get soon (probably by 8.09) is having the computer capable of being interactively interrogated in order to focus search questions quickly. (The big gain here comes from enabling the computer to ask the requestor for specific types of clarification in the disambiguation process.)

6. There have been, and continue to be, frequent large jumps in communications ('comms') technology beyond the telephone—which today means the cell phone (or, for remote field work, the satellite phone)—and radio, that now make it possible to reach anyone who does not block access (and many who think they *have* blocked access) almost instantly and almost without cost (Skype and its six imitators). They can be reached, not just with voice contact but with video, so important in serious efforts to benefit from face to face verbal exchanges. (i) This makes it possible for the use of online interviewing—and hence detailed case studies—to expand the face-to-face (F2F) accessible population as an alternative to traveling to visit the interviewees, at much lower cost, thereby facilitating cross-cultural representation (and, to some degree, expanding it to different linguistic groups with simultaneous translation). (ii) It also makes it possible for the evaluator to confirm, expand, and refine reports and accounts to a degree not previously possible and still not fully exploited. For example, in doing international aid evaluations, we will next year be able to send illustrated draft reports to country field staff for checking *and* round-table online video discussion, without appreciable delay in delivering the final report, and with a significant increase in validity. (iii) Webinar and many competing brands of software also make possible group meetings with consultants, and focus groups with participants, from all global locations, an almost certain way to improve critical dialog and information/perception collection, hence many evaluations. It also seems likely that evaluators will be able to create some useful improvements in Delphi technique by using new comms capabilities: for example, we might develop something we could call 'fast Delphi' with a spatially extended group, meeting sev-

[19] The National Center for Education Statistics is the most important location for these.

eral times within a day, to avoid the slow speed of traditional Delphi. (iv) Online surveys are of course now frequent and becoming a subspecialty with almost as much known about good and bad practices as in classical focus group methodology; upgrading these to video status will be a natural extension. (v) The addition of true 3D to the video presentations possible over the internet or via satellite will make truly competitive something that is currently only marginally effective—the use of distance consulting and interviewing and small group conferencing at a level of realism that will make it hard to justify real travel to meetings or consultations, especially in view of the carbon footprint dimension of evaluation. And, of even greater importance in personnel evaluation, the addition of 3D will make it possible to avoid the cost of travel (i.e., the handicap, for candidates, of not traveling) to interviews for jobs. It will also significantly enhance an approach that is already fairly functional—online teaching, covered in the next item in the Arabic-numbered list we are in at the moment. For example, teaching an evaluation approach like Appreciative Inquiry or Participatory Evaluation via text or lecture is, students report from analgous cases, *hugely* improved if high-definition life-size demonstrations by leading exponents can be added (with no-glasses 3D in 2010 and beyond). (vi) Looking once more into the near future, a reasonable bet is that the next large payoff from technology in the comms area is already in place in some of the large phone survey shops, or at least an approximation to it. This is the extension of the call waiting feature already available for domestic phone service, which now shows you a picture of the person calling as well as their name. The extension will provide every professional interviewer with a screen full of details about the person being interviewed in a survey, along with customized prompts for the next questions. (vii) Also on the list of comms-based technological jumps for evaluation, it is now becoming clear that we could come close to running 'shadow-conferences' in which at least the main speakers and a selection of others at the big professional conventions could be put online as simulcasts for access by those unable to attend. Since these annual meetings in evaluation are of great importance in keeping up with the field, it seems clear that making an effort to narrowcast a slice of them would show an appropriate concern to globalize evaluation, since we presumably believe this would be a benefit to those interested but unable to afford the trip. (viii) We are also approaching the era of 'Web 2.0' the mythical redesign of the Internet of which some tendrils are beginning to appear, e.g., 'cloud computing' (the use of online applications instead of software located in each computer). This will inevitably involve highly personalized sample selection, where the pattern of web and phone use by subjects will be used to select them. Integrated into cell phones as is already the case, and given that we are now approaching the point where 50% of the world's population has a cell phone, the web will clearly become the primary vehicle for survey work, including F2F phone interviewing.

7. It seems fair to say that we have now reached the point with online teaching where we can claim 'parity on balance' with onsite teaching, of evaluation in particular. That is, there are enough advantages to the online approach to offset the undeniable loss of some advantages pointing in the other direction, such as the greater inspiration of direct contact with the instructor, and with classmates, at least for many students.[20] Advantages for the online

---

[20] One problem with distance education for credit purposes has not been surmounted in any of the commercial packages although I have now developed a solution. This is the authentication problem a.k.a. control of the widespread problem of cheating by the use of surrogates to take the tests and

mode include: (i) access to instructors who cannot be afforded in person or who are unwilling to travel; (ii) access to the instructional process by students who cannot afford the time or travel costs, or physical difficulties, of attending a lecture in person; or who (iii) cannot schedule the times one is available in person; (iv) the ability to delay, replay, or slow down/speed up the presentation to suit the learner's preferences; (v) the ability to pull together a wider range of students, where diversity of occupation or income or age is important for the student and class discussions; and (vi) the ability for the presenter to edit and improve a presentation, piece by piece, until it become the best possible production by that presenter. This one use of educational technology today, i.e., 'e-learning' or remote instruction, which will be further enhanced—for many people—as we move into true 3D, means that we now have a genuine alternative to classroom and campus-based education, from a kindergarten (for some children at least) to the professional development domain. (A later section of this paper discusses the ramifications of this for instruction at all levels in all subjects.) Here we just stress that the globalization of evaluation is made possible by this single aspect of technology, since fast updating of remote countries' professionals is now possible in a way not previously done during decades of slow percolation.

8. Graphical information systems (GIS), three-dimensional global positioning systems (GPS/WAAS), and radio-frequency identification (RFID) are now available and beginning to be used in evaluation. They make spatial/geographical analysis possible which, in cases such as evaluations of crime-control programs, can be very helpful in locating success and failure areas. The third dimension can also be valuable in dealing with high-rise crime in housing developments, fire in office buildings, and rescue efforts in mountainous areas. Working in central Africa, in both coastal and central countries, the application that occurs to a technophile-evaluator is the control of illicit logging and poaching of game animals by using a combination of GPS, GIS, and RFID. Tagging a large random sample of trees and animals would make tracing their movements as easy as LoJack with cars in cities. It would be an easy application of a mid-priced technology with very high potential not just for arresting criminals but for its deterrence effects.

9. Almost everything mentioned so far depends on the computer, especially the personal computer. It's the tool that uses the software and for that matter is used to create it. The evaluator almost has to have one, and increasingly these days—especially if they do any field work—more than one. On this front, the latest breakthrough is the emergence of what is now being called the netbook, the ultraportable successor to the notebook computer. These were inspired by the quest for the $100 computer by the One Laptop Per Child project initiated by Negroponte at MIT, which culminated in a very advanced solar-powered small computer at around $200 (with reductions to come). These are now also available commercially for two to four hundred dollars from Acer (the Aspire), Dell (the

---

do the term papers. (The extent of this is often denied, but all the hard evidence available points the other way. In evaluating this situation, keep in mind that the use of ringers at the leading onsite universities (in the US, at least) also appears to be very extensive in large classes, so this is not a ground for condemning online college degrees as inferior.) The solution is a little complex, but it uses video and randomly timed online 'visits,' by Skype, to ask the students questions about their tests and papers with a government-verified photo on a split screen next to the direct feed on the instructor's desk.

Mini 9), and the first of the commercial entries, the Asus eeePC. They all have 9" screens, most have small RAM (magnetic memory) drives instead of hard disks, hence no moving parts to break down or wear out. They either omit Windows/Mac operating systems, keeping the price down by using an open access operating system and browser to access the Internet (hence 'netbook'), or the cheap Windows XP. With a weight under 3lbs (1.5kg), they provide the evaluator with a 'real' computer in the field with almost negligible portability problems. Current 'smartphones' also have keyboards but so small they can't be used extensively. Eventually, voice input will make the keyboard less important, but one can't—or does not always want to—speak out loud what one is inputting, and the smartphones aren't much good for spreadsheet work.

The future here lies with nanotechnology and it's clear that it's only a matter of time before the computer will be the size of a pack of cigarettes, or, using a thinner form factor, be capable of being stitched into a hat lining or a shirt collar, with the optional keyboard being printed on electrophoretic paper, folded to 3" x 5", and unrolled when needed. The step beyond that, already in the planning stage, means implanting the computer subcutaneously behind the ear and learning to type on a keyboard projected as a transparent film in the field of sight, injected into the optic nerve upon (mental) request. Hence the computer eventually becomes integrated with and simply an augmentation of the human brain, i.e., an artifact working to provide virtual enlargement of cranial capacity.

10. For everything that has been mentioned so far, including all the payoffs for evaluation, there is always the shadow hanging over them of possible breakdown, maintenance, and upgrading. And here many of the software applications mentioned can come to their own rescue—to some extent, at least. Not only do we now have full scale remote control of our computers so that we can allow the online technician to take over its operation and do the trouble-shooting directly, but when a part has to be installed—which can't be done remotely—or diagnosis requires actions outside the control of the computer, or the computer itself won't work at all, then the phone line and a laptop or PDA or mobile phone is increasingly able to transmit video of the problem machine or screen to the distant technician. So using these computerized enhancements to evaluation is not as risky as it used to be; not that it's 100% bulletproof.

11. Finally, we come to publications. It is still the case that scholarly discourse and research is much more influenced by what is published in scholarly journals and books than by what is online. Additions to these influential media are still largely inaccessible to most of the educated people in the world because neither they nor their libraries can afford more than a small sample of what is available and relevant. This is especially important to new disciplines like evaluation because the regular course instruction system, worldwide, is not yet geared up to provide evaluation skills and tools to the mainstream students. So for some time, access to these new tools, with their increased powers of resistance to economic exploitation and to inappropriate cultural globalization, was very much an elite privilege. This situation with respect to *current research and discussions* was in marked contrast to the situation with respect to knowledge at a more general and very slightly less current level, where Wikipedia and its imitators, and Google/Yahoo etc., have made access to much of the corpus available without charge. The traditional repository of such knowledge, exemplified by the Encyclopedia Britannica, now takes a back seat—not to be entirely dismissed, but clearly much less important because so out of date in the accelerating world of

research, and so expensive. Two new approaches, one from Google, hybridize Wikis with classical encyclopedias by using referees, experts, and volunteers, and some hybrid may eventually be the best option for the scholar.

The economic justification of the old system in the professional area has vanished, and could swiftly change much more. Technology has removed the cost barrier to access to evaluation, for example, by introducing high-speed copiers, online publication, and the commercially workable, 'publishing-on-demand' (POD) model, which almost completely eliminates inventory and overrun costs. However, social inertia, or vested interests, or lack of sensitivity to overseas needs in the professional associations, has delayed anything like the maximum possible helpful response. The difference between these associations is, however, more notable than the average situation. Towards the bad end of the scale we would have to list AERA (the American Educational Reseach Association), where the recent increase in membership fees has put them well over the average annual income in many countries ($120); their key publication in the methodological area, the third edition of ***Complementary Methods***, has recently been issued in hardbound at $175, paperback at $90; and access to their online discussion forums has now been restricted to members. In other words, a fairly complete closeout of most professionals in poorer nations from discussions and developments in educational research, an important lifeline to improving their own quality of education and standard of living. Not incidentally, this cost barrier discriminates sharply against students, despite some discounts for them, and against less well-funded college faculty and libraries in the US as well as overseas.

The AEA has done much better. Its online discussion forum is open to all, and it has supported a small effort at low cost book publishing. It is now easy to publish useful books in evaluation for sales in the low hundreds at about $12 per copy[21], but they are still usually made available at about five times that price, and sometimes ten times that price. Similarly, it is not too hard to make journals available free online with minor subsidies[22], but they are commonly only available at about the same multiple, either via expensive memberships or via high subscription or per-article charges. It is true that the low cost versions are not up to the highest commercial or university press standards in terms of copy-editing, but the content is widely thought to be comparable, on balance, in quality and utility. It seems clear that a thoughtful strategy would require most professional associations to use something like these options.

But there is a gap in the professional publications offerings, and it's an important one. We

---

[21] I produce the Monograph Series for the AEA at $16 (paperback edition), but that includes $4 for AEA publications funding (and $4 for the author). Choosing titles more for saleability instead of disciplinary importance, as most publishers do, one could do break-even publishing at $12. This is economically possible because the AEA announces the books, i.e., free advertising, and the editor donates his time as manager, editor in chief, etc., and no doubt others would be willing to do this.

[22] I established and publish the ***Journal of MultiDisciplinary Evaluation*** online, free, with the only subsidy being the very small cost to the Evaluation Center at Western Michigan of hosting the website (jmde.com). All labor is donated, a valuable experience for the students who provide most of it. This journal joins many others (now more than 1,000 refereed journals) in the 'open publishing' movement, and is further supported by very useful software funded by the Canadian government which is provided free to all free online refereed journals.

would greatly benefit from a library of mini-presentations by leading proponents of particular approaches to our subject—in our case, evaluation, but the point is equally valid in political education and woodturning—and mini-demonstrations by their choice of actors and subjects. Done for a big screen with HD and stereo sound, perhaps adding 3D in a couple of years, this library would make possible a huge jump in the effectiveness of teaching and training. Ideally, the professional associations should assume responsibility for assembling these, and managing them as a non-profit effort, with rental/sales fees for CDs (and perhaps online access via Amazon/Apple/Audible) just covering the administrative costs; grants and association membership fees covering start-up and add-ons.

12. Finally, we come to a technological phenomenon that is both important and almost evaluation-specific. I'll call this public participatory evaluation (PPE), and it was never predicted; indeed, when it appeared, its failure was predicted with an equal lack of validity. The term refers to a range of phenomena exemplified by Wikipedia (and its many imitators) and the online feedback ratings offered by many large retailer and news services (e.g., Amazon, BBC, CNet, TravelMole). It involves three factors: (i) near-zero supervision; (ii) online reporting, often on a highly publicized website with million-plus daily hits. There are now many websites that were set up for, and are devoted entirely to, this kind of 'amateur evaluation'; (iii) dealing with both controversial factual issues and estimates of merit. The recent highly credible run-off between Wikipedia and the Encyclopedia Britannica[23], which resulted in a tie, is a fair indicator of the general level of success of PPE. Established evaluation sources, e.g., Consumers Union, have been slow to respond to PPE, with consequent loss of credibility and utility; and the relevant evaluation associations will need to respond or risk the same consequences. And PPE has a wider political range than conventional evaluation, as is discussed in the next section.

The preceding list is not claimed to be complete, but it is perhaps enough to indicate that, taken in conjunction with what was said earlier, and is said below about online teaching, technology has produced a huge accelerant to the internationalization of evaluation if not globalization in its less attractive aspects.

The main conclusion for our purposes in this essay is this: technology is as crucial to evaluation as it is to education, and hence doubly crucial to educational evaluation. Understanding its importance requires understanding technology, and understanding its potentiality for evaluation is arguably a professional duty in evaluation and indeed in any profession. In particular, understanding its potential importance for giving the world a high level of expertise in evaluation is a mirror of the main theme of the paper—a case of largely unfulfilled potential. Our country has done marvels in donating food to help meet global needs; it has done considerable damage in blundering efforts to give them technology. The reforms due to the appropriate technology movement have corrected much of that tendency;[24] but the errors of omission now need to be addressed, and using technology to improve access to professional evaluation materials should be a top priority.

---

[23] This competition was set up and run by ***Nature*** and a typical report is at Silicon.com, dated 12.16.05. This, and other reports on the same event, are on Google (search on <Wikipedia vs. Encyclopaedia Britannica>).

[24] See Wikipedia and thefarm.org .

5. Technology in Education: (II) Media and Presentation

When educators today talk about technology in education, as pointed out earlier, they are usually referring to computers. But of course the effect of technology on education was enormous for thousands of years before computers arrived. How did this all begin? As pointed out earlier, it began with the mightiest of all technologies, the first general purpose artifact, the most powerful lever in the development of all that came later: language itself. Of course at first this was only the spoken language, but then, unknown millennia later, it gave birth to its mightiest offspring—written, painted, or incised language. Language was by no means the first technology; it seems clear that our prehuman ancestors, like our close relatives the chimpanzees, made tools and shelters before they could talk about them in anything deserving the name of a language; and they appear to have had domestic fire well before language. But language was the artifact that gave us the edge over all other species.

Now the technology of language was not necessary for education, since teaching by example, perhaps even teaching assisted by standardized gestures, precedes teaching by linguistic communication and is common in many species. But language was the great accelerant of education and became the cement of organized education, of schooling as we know it, in the way that communications—by courier, post, heliograph, telegraph, or internet—are the cement of government. So when you hear teachers speaking of technology but meaning just computers, remember that their very act of speaking illustrates a technology a million or so years older than computers, one so deeply ingrained in our lives that we don't even realize that it, too, is a technology. And the failure that is evidenced by that careless way of speaking is a constant reminder of our deep failure to understand technology and its role in our lives.

This discussion of the roots of educational technology serves two purposes. First, it's a useful introduction to the evaluation of the current and recent situation that is the main purpose of this section. Second, it's part of the demonstration that there really is substantial content for a curriculum in technology education that we have been trying to justify as an important component in public education.

Students are sometimes resistant to the idea of language as a technology, probably because it's so different from the other things they are used to identifying as artifacts and hence part of technology. In convincing them on this point, the computer itself provides some support. While for the first million years or so of the history of technology, most of it was hardware, the computer has brought us back to the realization that those things don't run without instructions, and the programming languages needed to convey those instructions are clearly artifacts. So we are now perhaps more willing to accept the idea that the language *we* use to 'run' each other had to be created, and its creation has to be recognized as one of the greatest artifacts of all. It's as well to remember, too, that there wasn't just one language for all people; there were many, just as there are many alternatives to almost every invention that comes along. Worth remembering, too, that every living language is still evolving, still changing in many ways that illustrate creativity and adaptability; that's also a key property of great technologies, not just computers but cars and knives and fish-hooks.

Now the 'written' language, to use the conventional term, although its early forms were not much like writing as we use the term today, could hardly exist without another technology—tools for writing. At first this may have been the hand, but soon the smashed end of a stick was used as a brush to paint on cave walls, or the sharp end of a stick used to write in sand; later it became the stone mason's chisel, to write on rocks that would last for thousands of years. The tools shaped the kind of artifact that could be used for letters, and obviously the medium available did so too. A flexibly written language, usable on multiple media, by contrast with pictographs and proto-writing, is only about six thousand years old, a late-comer in the history of technology but another great accelerant, one of the elements that made distant communications possible without a person having to make the trip.

The next great step was printing, and then the challenge was to make communications possible without a physical record having to make the trip: in other words, signal communications. Flag signals ran the great fleets of early empires, smoke signals worked fine for indigenes in North America, fires on the mountain tops worked fine in Europe, and the heliograph worked fine on the Northwest Frontier in India. But for the full riches of existing languages in the signals, and more privacy, then the key invention was encrypted Morse code over the airwaves in the 19th Century, leading to emailing and texting today.

The assistive role of writing technology manifests those technologies as books and letters by messenger, i.e., 'distance education' for to the rich or powerful a few thousand years ago, and then the correspondence schools of the early twentieth century, the radio for the School of the Air in outback Australia by 1948, and online courses today. On the proximate scale, the assistive use of projectors, copiers, and printers has been useful if not revolutionary; but the use of Braille and then OCR (optical character recognition) was revolutionary for one group of handicapped students, where lip-reading, sign language and voice into text conversion served another. These were all great developments in educational technology, opening doors that were completely closed, or making the opening easier. And of course, as with the development of wheelchairs, it was not only the doors to classrooms that were opened but also, at least to some extent, those to jobs and entertainment.

The earliest form of these developments did something else important besides producing the results so far mentioned: they made possible art as we know it. In fact, the art preceded the written language, in the cave paintings at least, and also in ornaments, clothing, and tool design. Divisions of art such as literature and theatre, dance and sculpture, music perhaps more than any, all depend in their own special ways on technology to exist, and developments within technology for their own development. While poetry is not thereby shown to *be* technology, in its written form it is entirely *dependent* on technology. And the writer in a discipline who invents a new term is well called a wordsmith, for a neologism is technologically dependent on writing, and its adoption is an instance of technological dissemination very like the dissemination of a better axe or sword or shield.

The study of language itself, at least from a certain point of view—not likely to be well received by the English Department faculty or linguists—thus becomes a branch of technology studies. The search for the laws that govern it—the rules of grammar—then becomes a kind of science, linguistics, remotely analogous to the science that emerged from the search to find the laws governing the passage of rays of light through a prism, the artifact that helped Newton found a branch of physics. Of course, it's more like a social science that a physical science, i.e., a science about human creations rather than natural phenomena; although it's also a relative of (pre)history since it's a study of a human creation of the long past. So the study of grammar, surprisingly, thus becomes one of the more serious curriculum components devoted to the study of a technology. Like all such studies, its value depends very much on the extent to which it assists us in a practical enterprise, and there its recent value is not clear.

Some of the implications of seeing language as an artifact are interesting. For example, it's well known that children can pick up languages much more readily than their elders, and this may be related to the fact that young people master technologies more readily, because understanding them is understanding something very like a language, complex, subtle—and functional, but for older people, often opaque to a degree that the young find mysterious. This suggests that an overlooked sales (and employment) niche could be populated by hiring elderly people rather than young geeks to write the manuals for cell phones and automobile navigation systems.

A more important possibility, from the educational point of view, is that one should write technology curricula for K–12 in terms of devices that the young understand and enjoy fiddling with, rather than in terms of the mathematics and science that grown-ups develop to explain the underlying physics of artifacts.[25] It seems to make some sense to think that there is a kind of 'language of artifacts ' that some people, particularly some young people, acquire very easily, and that they consequently find it easier to get artifacts to 'talk to them.' It is certainly not the language of science and mathematics and we do a great disservice to young people by ignoring this gift, and to society by ignoring these 'linguists' and their talent.[26] Dislike for this kind of thinking is widespread; for example, in the final draft of the US 'national curriculum for the 21st century in science and technology,' which I was asked to review, computer science was not listed as a science. This is a disgraceful and dysfunctional example of 'science-ed elitism,' and one more example of the extent of the failure of current education to provide appropriate support for a technology-dependent and technologist-dependent society.

It will be obvious that a parallel line of thought shows how technology was instrumental in the development of mathematics. While counting no doubt preceded mathematics, arithmetic—and of course geometry and further mathematical developments—required some

[25] When asked to do this by the Ministry of Education in Western Australia, I built an experimental technology curriculum around a suitcase full of interesting artifacts, e.g., five types of can opener, four-bladed and other non-standard boomerangs as well as the conventional kind, and got the students engaged in comparative evaluation of them. It was extremely rewarding to hear that in the field trials some of the youngsters on the aboriginal reservations who had not spoken a word for the first few months of the school year had become entirely engrossed in discussions of this activity.

[26] The present (2006 figures from NSF) on the disastrous collapse of female enrolments into computer science college courses and degree programs is not due to any feminine lack of ability with mathematics and hence with computer languages. This is shown by the results from countries that have leveled the playing field in mathematics and science instruction, notably Australia (OECD figures), where it is now the case that girls outscore boys in all school subjects. At the college level it is more likely due to the bad design of the advanced placement course in computer science, which treats it as a branch of mathematics when it's at least half a branch of technology and should probably be taught as a hybrid.

written record, probably a refinement of the primitive marking systems used for early husbandry, then for bookkeeping, and eventually taxation in Egypt and China.

The final implication of this line of thought to be mentioned here is, as may have already occurred to the thoughtful reader, that these early technologies—writing and its refinements, with a little assistance from measurement devices—were what made science possible. If the father of science was the inquiring mind, looking for explanations and predictions, the mother was technology, seeking to improve practical devices.[27] So the ultimate irony about the idea of technology as applied science is not just its historical absurdity, but its logistical impossibility. Mere speculation about the nature of the stars or the origin of the universe or the cause of floods is the intellectual precursor of science, but *science itself* required records, written calculations, and tested predictions.

As these disciplines developed, in the temples and the counting houses, education about their practice went along with them, so that there would be successors for the roles of expert users of the new sciences and mathematics. Thus the early history of science and mathematics education, as for science and mathematics themselves, is inextricably linked with the middle history of technology.

But none of these assistive technologies, however vital, were *pedagogically revolutionary*. That is, to oversimplify somewhat, none of them altered the basic modes of teaching, only the machinery of delivery. Especially for those suffering from physical impairments, these were huge breakthroughs. However, technology *can* assist in a pedagogical breakthrough, one that produces a huge improvement in the amount and/or quality of learning for a much larger group that is badly served at the moment, and indeed considerable benefits for virtually every student. And that brings us to the last chapter of our story.

It's time to pull all of this together. The interactions between Education, Technology, Evaluation, and Globalization would take many books to deal with adequately, but the aim here is to call attention to one particular strand that seems particularly significant. There is not space here to document in full detail every data point on the curve I am describing as the "decline of the best," so my perception of the events described below should just be taken as the view of a close, critical, participant observer, and the overall view as a possible interpretation; even that modest status should make it something to be considered by the thoughtful reader, which is all I hope to achieve.

How can technology radically improve pedagogy? First, we need to identify the direction in which improvement lies. We have a very long tradition of one ideal in education that is worth some detailing. It begins with Socrates and the Sophists some two thousand years ago. Though the former had great disdain for the latter, certain features of the practices of each have long stayed with us as components in a widely respected educational process. From Socrates' method we garner the idea of 'non-directive teaching' that is, a non-didactic highly interactive method aimed at bringing the answers out from the learner via intelligent interrogation. From the Sophists, who made their living by teaching the skills of argumentation (which Socrates thought rather despicable), the great skill was the ability to bring logic to bear on specific practical issues. They would undertake to compete with ex-

[27] Similarly, in the arts, songs and sagas were the fathers of poetry and fiction, but the scrip and the scroll were the mothers.

perts in substantial subject matter fields and beat them in debate—the winner to be decided by the audience—through using their better powers of reasoning. There are at least these elements in the combination of these approaches:

• A high degree of individualization. (In both approaches)

• Intelligent questioning aimed at testing and encouraging deep insights rather than long lecturing on deep subjects. (Both)

• Immediate feedback as to whether the answers being given are correct. (Socrates particularly)

• Showing payoff in areas of recognized knowledge and expertise (Sophists particularly, but also Socrates with, for example, the slave boy)

At least some of these values are also highly respected in certain Eastern traditions, e.g., in the teaching style of the Buddha, and in the elder traditions of many indigenous groups. They are also closely connected with the kind of teaching that the tutors hired by wealthy families in the East and West expected from these tutors in the times of the Persian and later empires down to the British Empire, since the tradition was often not to use public schools even if available. When the children grew old enough to go on to professional training, we find them attending the famous tutorials required of undergraduates every week at Oxford and Cambridge, or the tough questioning on case studies in larger groups in the better law schools, and in their clinical training in medical schools.

Clearly, cost considerations make a true tutorial approach hard to afford for all students, and a weekly meeting is a poor substitute for the daily sessions of the Greek Academy. The usual large introductory level university lecture course, with as many as several hundred in the audience, is a long way from the ideal, but is usually supplemented by 'tutorial sections' with a score or so students in each, where slightly more individualized attention is possible. However, although all of these can be questioned at one time, only one answer can be followed up with feedback tailored to it, and with another question tailored to that.

If all members of these smaller and larger groups in classrooms were similar in their learning speed for this subject, at least the rate of progress even if not the individual answers could be matched by the teacher's rate of presentation. If their prior knowledge of the subject under discussion is closely similar at the beginning of the lesson, at least this foundation can be correctly assumed in the questions and comments produced by the instructor. If neither of these conditions are met, and they are almost never met, virtually all analogy to a Socratic dialog vanishes. And the more varied the learning speed and prior knowledge, the more difficult the task of instruction will be. The current US norm in the K–12 classrooms, which is age-graded membership, often with high attrition rates, is so far from the ideal that it is, in the view of many teachers and teacher trainers, very nearly hopeless.

The main reason for setting up such a difficult teaching/learning situation is an oversimplified conception of what fairness requires, combined with an apparent inability to enforce and support better approximations to a genuinely fair learning environment for children of all ethnic and economic backgrounds. (A genuinely fair environment is one that maximizes the learning opportunity for every student, of whatever knowledge background and learning speed.)

Now, what can technology do towards solving this problem? We have a long history of experiments in this direction. I will briefly review some case studies that I have personally observed and examined.

Case A: B. F. Skinner and his teaching machines

First, a word of warning. I'm going to be talking about the technology of teaching machines, not the learning theory of their inventor.[28] If you have objections to that theory, as I do, don't let it stand in the way of evaluating the technology, any more than you let Newton's eccentric theology interfere with your appreciation of his invention of the Newtonian telescope.

In the course of working on a long and highly critical study of his theories of human behavior in the 1950s,[29] I became good friends with the psychologist B. F. Skinner, and he invited me to visit him at Harvard in order to look at the way he was teaching his introductory psychology course, using teaching machines. These simple devices were designed to implement his theory of learning, which is committed to the idea of presenting material in small slices, asking questions about the student's understanding of it immediately after every presentation (a few times a minute), and providing the correct answer immediately the student's answer was completed. The text material was in the form of a roll of paper, loaded on two rollers in a box with a small window in the top where a short passage could be read. A question followed each passage, and a slot allowed the student to write in an answer onto the scroll; advancing the scroll further revealed the answer to the question. The handle which advanced the scroll was governed by a ratchet that prevented the student from going forward to look at the answers before writing in his or her answers, or going back to check the relevant material before answering the question. Each student advanced at his or her own pace, which varied considerably. Progress was measured by regular tests ("quizzes"), given at fixed points in the material, for which the answers were written by the student on a separate sheet that was collected and scored by the teaching assistants, and returned quickly.

The uniquely Skinnerian contribution, apart from his typical connection of theory to practice—by inventing a technology to incorporate it—was to write the material so carefully and at such length, and rewrite it so often in the light of its success, that the students made very few mistakes, hence nearly always got the right answer, i.e., they received almost entirely 'positive reinforcement' in his terminology. These rewards for successful learning consisted in seeing the correct answer, which motivated them to continue and to try to succeed again. Note that this approach corresponds roughly with the tenets of the Socratic method listed earlier, and the whole approach matches the Sophist model. This arrangement was reported as producing a major increase in the scores of students on the same tests that had been used previously, when a more conventional lecture-based approach was em-

---

[28] Skinner was always careful to point out that he had a precursor, although of a very different device (self-correcting tests). Unfortunately, none of the standard references mention this so I can't give the appropriate credit here.

[29] "A Study of Radical Behaviorism" in ***Psychology, Psychoanalysis and the Foundations of Science (Minnesota Studies in the Philosophy of Science)*** , Vol.1, 1956, eds. H. Feigl and M. Scriven, University of Minnesota Press, pp. 88-130

ployed.

The machines occasionally broke down; if that occurred, the students would take the roll out of the box and write answers directly on it. Of course, they could then 'cheat,' by looking at the answers before giving their own. The interesting result was that they then did no better or worse on the check-up quizzes. Capitalizing on this, in typical Skinner style, it was tested as the standard procedure, with the materials in book form (much cheaper than the machines) and the answers only a page or two away in case the student wanted to 'cheat.' Of course, they could not cheat on the check-up quizzes, and had to retake the chapter if they did not pass them on the first try (a little negative reinforcement, occurring rarely once the materials were fully developed).

Case B: A. Calvin and programmed texts

The approach via the book version was seized upon as a new way to make texts more effective—they were called 'programmed texts' (PTs)—and many texts in many subjects were written in this way, for K-12 as well as college courses. They were completely individualized in terms of the student setting his/her own pace, and this at least partially, and sometimes more than completely, offset the huge compromises of the mixed-skill mixed-background classroom.

I visited the headquarters of one of the organized efforts to do this, led by Allen Calvin at Hollins College in Virginia. I watched an eighth grade class in the local school using a programmed text for ninth grade algebra. They were supervised by a teacher who was prohibited from giving any assistance other than to keep order (to avoid the possibility that any improvement in the results was due to their teaching), and they did better than the ninth grade students under the usual regime. Of course, with a teacher using a PT in the normal course of events, who would be encouraged to give individual assistance to those who were having any difficulty, it would be reasonable to expect even better results, and presumably even more so if the students were in fact the ninth grade students normally thought to be intellectually more easily able to handle algebraic notions. Helping individual students produces tangible results is for many teachers more gratifying than trying to get all of a large, heterogeneous class moving through a textbook written for one level of learning ability, or through lessons at one level. But even with these handicaps on the 8^{th} grade teacher, the PT beat the 9^{th} grade teacher.

Bill Sullivan produced some excellent PTs funded by the (then so-named) Bureau for Education of the Handicapped, aimed at slow learners, which achieved the seemingly paradoxical goal of teaching reading. (Of course, they did this by using pictures to begin with.) These were notable for two reasons: (i) they were also very good for teaching average students to read, which illustrated a general point about PTs—fast learners simply moved faster through them, they didn't need different PTs; (ii) they were run in 'horse races' against the usual commercial texts with experienced teachers using them (at Cherry Creek in Colorado, and in Philadephia city) and won handily. So PTs not only seemed good but proved good.

PTs had three real drawbacks and three 'artificial' problems. To begin with the real problems. (i) They were costly, and hard, to *create*, if made to high standards, because—however good the writer—it took many rewrites to reach the required standard of about

85% correct answers on the embedded questions and a high score on the check-up quizzes. Each draft had to go out to a group of students at the target level and a teacher, all of whom had to be paid—as did the writer—and then the material preceding any questions that were not meeting the correct-answer standard required had to be rewritten.

(ii) By the time they did reach that standard, they were very bulky—hence, costly to *manufacture*. (iii) If answers were written into the blanks in the PT, the book was rendered useless for another student, unlike many texts, since they can usually be reused; another expense drawback. The second and third problems are eliminated if the materials are presented online. The first is greatly reduced if the trial students are online, too.

The 'artificial' (a.k.a. possibly fixable) pseudo-problems were these. (iv) Students finish a PT book at different times, and so the whole (completely artificial) cohort system of grade levels does not work in the usual way. The students didn't all 'finish' at the same time. This upsets administrators, and parents, who are used to the standard system. It's easily handled, in the way common overseas, by making class groupings achievement-based, so that a student moves up to work with another group in that subject, as soon as the latter group completes the text. This may be in week three or thirteen of the school year. This approach has the added advantage that it destroys the (bad) idea that students 'belong' in a single grade level for all subjects, which in turn undercuts the main basis for the 'dummy' stereotype. The problem of finding your age-mates as playmates is self-solving in the playground or at social events for age groups if you really think this is important (most kids don't), or if not, there are many ingenious suggestions in teacher guides for the UK.

(v) Some public relations genius looked at the Hollins-type field trials and saw that PTs worked without teachers—and having heard incontrovertible evidence that some teachers somewhere could not pass the state exams for students in the subjects they were teaching, coined the term 'teacher-proof' for PTs. Immediately, since this was a time of teacher surpluses, the unions got nervous and teachers began to see PTs as a threat to job security. Of course, as indicated, there's no reduction in the number or need for teachers, meaning good teachers; in fact, the PTs are a path to better quality of worklife for teachers.

(vi) Greedy publishers saw that they could put out a pseudo-PT by just taking an existing text, and reformatting it into the numbered paragraphs (called 'frames' in a PT) which contain the embedded questions, adding some embedded questions, and skipping the field tests and all the rewrites to meet success-level standards. The results were so bad that there were actual student revolts against using them, and the public, absent adequate guidance from schools, education departments, the feds, and curriculum experts, was unable to distinguish the good PTs from the bad ones and voted against their use. Gresham's Law strikes again. The market for PTs evaporated. They could have been saved, but it would have taken a deep understanding of technology as well as PTs to mount the effort, and few people met those requirements. As it was, the last best hope for saving the slow learners and the fast learners went with them. It was another step in the decline of the best hope.

Case C: O. K. Moore and the 'Talking Typewriter'

Around 1960, I visited a kindergarten in New Haven where a very different application of something like Skinnerian principles was operating. O. K. Moore, a faculty member at Yale, was running a 'talking typewriter.' This consisted of a mainframe computer (there were no other kinds in 1960) hooked up to a tape recorder and a big monitor plus a keyboard, and was used to teach the children touch-typing, reading, composition, vocabulary, and spelling. The kids took turns between play activities of the usual kind and occasional sessions on the talking typewriter (lots of kids for one typewriter), which was programmed with the usual instructional materials (with voice output) for learning touch typing, plus they could compose stories and get their spelling and some grammar checked. They never *had* to use it, always having the choice of continuing with their other play. This approach worked well, and I've never seen any evaluation pointing to significant downside, other than cost ($35,000 in 1960, about $500 in 2008). Moore's generic plan was to make virtually all learning a playful activity: see more on this approach below. Notice the way in which the Socratic principles are implemented in this case.

Case D. Fred Keller and the Personalized System of Instruction (PSI)

At the post-secondary education level, the Skinnerian banner was carried by Fred Keller, a psychology professor in Arizona, whose PSI movement had a big spike of success in the 1970s. It used more conventional text materials, but, following Skinner, dropped most of the lectures, allowed self-pacing with mastery-learning required tests, immediately graded by TAs. It was very successful for a while, although, a big cost here as with PTs and the teaching machines, it required a great deal more work from the instructor to prepare more and better materials and tests. The Google coverage on this is very good, and the first entry (as of 4.21.7) by Howard Gallup[30] is an excellent treatment of ambience as well as details and problems. The student ratings are nearly always very high *if* the program is implemented conscientiously; but after 1980 the implementations fell away. Now we are beginning to see why these improvements are low on sustainability—the load on the instructor is high. But we can also see the way to improve that aspect of them; more on this under the next case study.

Case E: The Highly Interactive Paradigm (HIP) in Higher Education

In the last decade or two, this approach (for which the above acronym is not widely used), has emerged in its own image but very much in the tradition of Cases A through D. The explicit ties to Skinner, often ignored in discussing Case D although implicit nonetheless, are not present and the implicit ones more tenuous; the technology is different and more conspicuous; the concessions to tradition slightly more extensive. But the thread back to Socrates and the Sophists is perfectly clear. In perhaps its strictest and most successful form, the key elements are:

The 'lectures' themselves occur on a regular schedule but are highly interactive in a new way, described below.

The feedback is very fast although only partly immediate (the part that occurs during the lectures) and is not fully individualized.

[30] http://www.google.com/search?q=Keller%20plan&ie=UTF-8&oe=UTF-8

There is heavy emphasis on deep insights instead of either rote memory or routine application of rules.

The results of full implementation are not only very large improvements in score levels over traditional lecture-centered approaches but also show improvements in quality as well as quantity, and are more decisively demonstrable than in any other model.[31]

Again, the increased time costs for the instructor are substantial, as is also the case for the students. The increased time is mostly in materials preparation, and more testing/grading, and that of course can be covered by groups joining to develop the materials, as has happened with the NSF-funded 'calculus-based first physics course' developed at Georgia Tech, Purdue, and North Carolina State.

Unfortunately, a new cost to the instructor is added: the student ratings tend to go down, rather than up as with the Keller plan. This is almost certainly due to the students' reaction to the increased workload, given that they are not in a position to judge the improvement in the quality or quantity of learning (since they only take the course once). They probably also miss the showy quality of entertaining lectures with good movies and graphics—but that is small consolation to faculty who are being judged by systems of teaching evaluation that depend heavily on the student ratings, as many do.

The novelty about the 'lectures' includes three notable features (in the full model): (i) they consist mainly of question-asking, which begins with a chance for students to ask the instructor questions about the assignment they have just completed and turned in at the door; (ii) and then turns into a session, occupying most of the class time, of questions from the instructor, testing comprehension of what has just been explained or said, or tested, using a multiple-choice format but typically with a large number of options , e.g., ten, to reduce the success of guessing; (iii) their answers are recorded by students using 'clickers' which are hand-held devices the size of a pack of cigarettes with several buttons on them, which transmit by infra-red or radio to a cumulative display at the lectern, showing totals but not seat origins; (iv) the questions asked of the students are typically based on the results of the homework from the latest assignment, which is handed in at the door coming in, and scored by a scanner with the results going to the instructor within a few minutes of the start of class (and to the students at the end of class) so that he can see where extra explanations and questions are required.

There are also prepared resources of 'clicker questions' tied to the general topics being addressed, on which the instructor can fall back if no strong needs emerge from the gaps in understanding displayed in the homework results. Of course, the teacher will often want to expand on, illustrate, or modify something in the text or other materials s/he is using; and can compose questions, on the fly or in advance, on these addenda and corrigenda.

No-one familiar with the history of educational reform will be likely to rate the chances of survival of the HIP model as very high. The academy, like most organizations, resents changes. Despite the frequent complaints of left-wing leanings by faculty, they are no more

[31] Excellent details on this approach, including the decisive evidence of its success, can be obtained from the web guardian of these archives; a search on "Richard Hake" (with the quotes) in Google Scholar, for example, will find them.

sympathetic to reform aimed at 'citizen rights' i.e., student rights and benefits than big companies, bureaucracies ,and conservative legislatures. It's pretty clear that the main reason they resent changes is that changes mean work (mental and physical), not a belief that the status quo is in fact superior, a case which simply cannot be made in the HIP case. Deans and provosts and state and federal ministries of education should be leading the way to support change with the HIP credentials; instead, they all too often prefer to side with the general faculty inertia and faculty unions rather than struggle for improvement. Not too surprising, perhaps, given that their long-term plans usually include retiring from administrative chores back to a faculty position. But let's be clear; backing down on this kind of breakthrough is simply one more case of the decline of the best.

The HIP approach does not match instruction to individual needs very closely, compared to the Skinner teaching machine or the programmed text, or the PSI approach, but it does focus on the needs of the group of students, as evidenced by their success or failure in answering the frequent quizzes, and the questions they raise in class. So it's a great step forward from lecturing combined with infrequent tests. It might be called a groupalizing approach rather than an individualizing approach.

Now, after one last case study, we can try to broaden the field of application to the global arena.

Case F: Alfred Bork and Global Computer-Assisted Instruction (CAI)

We have not provided earlier case studies of CAI (the abbreviation I prefer, for its literal meaning), partly because the problem of generalization is often very difficult, and partly because I wanted to first build a defense against the serious negative criticisms of CAI from Larry Cuban[32]. It seemed more valuable, since we are here looking for ways for technology to produce breakthroughs in pedagogy, to first establish the research basis for the learning principles and evidence for their practical validity. Once these are clear, they can then be incorporated fairly easily into software that will produce good results, provided that the subject-matter content is of the highest quality. Professor Bork of the physics and computer science departments at the University of California/Irvine, put his introductory physics course online with NSF funding and considerable success in the 1970s, and then became interested in the wider applications of this approach. In the book, ***Tutorial Distance Learning***, (Springer, 2001), he and his co-author, Sigrun Gunnarsdottir, make a case for what we need as a major technological breakthrough in education.

The case has essentially three well-supported premises: (i) the present system of basic education in the US and many other wealthy countries is graduating or failing out a large number of illiterate and/or severely undereducated students; (ii) the evidence supports the idea that good education requires, or at least is well served by, highly interactive, highly individualized, instruction of high content quality matched to high needs; (iii) the system for preparing teachers in the US is doing a bad job of quality control, since many teachers,

[32] I am quite familiar with CAI's capabilities and limitations, since I did the most extensive evaluation of the largest, and best-known, CAI site, the PLATO system installation at the University of Western Australia, as well as the evaluation for NSF of Alfred Bork's CAI system at the University of California/.

especially in mathematics and science, are demonstrably not competent enough in those subjects to answer either the standard test questions or many questions from their students. (iv) These teachers are also faced with the extremely difficult task of dealing with classes exhibiting highly diverse abilities/achievements. Therefore, the argument goes, the best solution is to use computer-tutors for which the materials are written by the best in the business, that are programmed for individualized and frequent question-asking with immediate feedback of correct answers, with help for erroneous answers, and which are backed up by teachers who help where they can, and maintain order and interest in various ways.

This argument appears plausible, given the evidence at the moment; and its global extension is compelling. It begins with some further powerful and well-supported premises: (v) the budgets of most countries in the world cannot, now or in the foreseeable future, possibly cover the creation *and maintenance* of schools staffed by trained teachers operating with even the largest student/teacher ratio now considered workable; (vi) it seems very plausible that a population with a high level of literacy is essential for a secure democracy; (vii) it is clear that these countries, with only a small amount of start-up funding from the usual donors, definitely can afford to buy and maintain computer-tutors with voice input and voice output, programmed entirely in the indigenous language. The MIT project has come through with the hardware end of this; powerful solar-powered computers at $150 each, this price probably to be reduced with more orders. Only the software needs to be funded, for a few schools in one or two countries, as a trial on the results of which further expansion would depend.

For more than a decade, Professor Bork has tried every source to get funding of the modest first trial of this proposal, with excellent supporting letters and evidence, but with a complete lack of success.

Part of the explanation of this situation, which appears bizarre on the merits, is contained in an answer given by Bill Gates at a news conference recently. He was asked why a billionaire whose fortune is based on computers, and with a charitable foundation doing wonderful work overseas, is not supporting computer education there. With commendable cultural sensitivity, he replied that we must be careful not to export solutions that have worked well for us here, but might not be appropriate elsewhere. Unfortunately, this is rather like the president of the American Medical Association referring to the medical system in this country as “having worked well for us.” If ‘us’ means the doctors, or the people in the top decile of the income distribution, it’s true they’re doing very well; but that’s not what 24–50 million uninsured people in this country think, along with a good many supposedly insured people who have been or will be refused payment for vital conditions on some quibble or another. Computers as *tools*, outside education, have worked quite well, although not as well as was supposed, as we found out when careful ROI (return on investment) studies were done. These uncovered the true size of the support costs, which showed that, as then being installed, computers were not paying off on balance across all installations, although when carefully installed to match need and abilities to use them appropriately, they pay off very well.

But the computers as *teachers*, in particular as teachers located so as to help solve the extremely serious crisis in US education, have apparently not served us well, although the case studies listed here suggest they could have done so very well. And that conclusion has apparently led to the refusal to fund Bork, and Gates' skepticism about globalizing them in classrooms.

Which brings us to Larry Cuban's good book on the failure of computers in the classrooms, ***Oversold and Underused: Computers in the Classroom*** (Harvard, 2003). He makes the case for failure very well, though it's worth looking at a careful review that draws almost the opposite conclusion, done by the Northwest Labs in 1991.[33] By and large, the situation is that the computers have not done a good teaching job on the standard curriculum, though they do a fine job as a tool for word processing, web scanning, income tax preparation, etc., and are not too bad, though not terrific, at teaching how to do those things.

Since I have just spent many pages arguing that several successful ways to reform teaching of conventional curriculum could be further improved by computerization, how can I agree with Cuban? The answer is that the good ideas were not computerized, and the computerization that was done was not based on the good ideas, but upon drill and practice or other routine, text-like, rule-based approaches. As I wrote the first draft of this, at the end of April 2007, a new government study has just been released that shows the time spent by students on computers teaching mathematics has not paid off. Surely more support for Cuban's skepticism about the 'computer solution'? Not at all; it's just more evidence of his mistake. Computers are not magic, they are just technology. Kids weren't mastering mathematics by reading textbooks, either; textbooks are just technology. And 'programmed texts' are just technology; the greedy publishers who wrecked that line of progress showed us that the magic isn't in the PT format, any more than it's in the book format. It's in the *optimization of the medium*, whether the medium is print, the PT format, computers, or computer software. The key issue is what *can be done with the medium* not what the medium does by itself. Technological breakthroughs are about what can be done with the invention, not with the invention itself; and what can be done takes a tremendous amount of work to get done.

The problem that led to the Cuban verdict thus comes from a profound error about technology. If a physicist has a bright idea, he sends it to a journal, and a couple of reviewers have a look at it; if it looks promising, they vote to put it in the journal and then if it's really any good, it will be cited and discussed. Bright ideas just 'bubble up' through this easy path to the sunlit surface of professional and perhaps even public recognition, as with Einstein's theory of special relativity. If technology is just applied science, as some dictionaries would have it, that model should be more or less the story with good ideas in technology. Since computers in education haven't 'bubbled up' by now, they just can't have 'the right stuff.'

But that's all just a fantasy as far as technology is concerned. That's not how technology reaches the sunlight of public acceptance. A bright idea in technology is like a larval form laid in the mud and boulders on the bottom of a turbulent stream. It has to *struggle up* to get through the mud and the jumble of boulders; then it has to *fend off* or by good fortune

---

33 http://www.google.com/search?q='computer-assisted%20instruction%22&ie=UTF-8&oe=UTF-8. It should also be noted that more than a million students in K-12 classrooms are using computer-assisted instruction in 2008 (several reports in the usual search engines), most of them getting good administrative reviews, probably because of the great advantage they have in facilitating highly individualized rates of progress in the usual mixed-skill classrooms.

avoid the predators, fish or fowl, that think it's lunch; after that it can still get caught up in the current and *swept away* downstream to the salt marsh which will kill it. Only if it avoids all these fates, then and only then, does it even get a chance at swimming hard to get to the surface. It has to work all the way, and mostly it won't get there, even if was every bit as healthy as, or indeed much more healthy than, the ones that do. It needs help to do as well as its merits justify.

So the evaluator of educational technology has a very different job to do from the reviewer of ideas for inclusion in a scientific journal. He or she has to identify the meritorious larvae *and the prospective barriers to their legitimate success* and assess the possibilities of ways around these barriers, and then make all of this clear to all the parties that can do something to ensure success for the best. That's what Cuban misses. He does see the problems that have prevented CAI from helping many millions of kids and adults, here and overseas, the problems being: (i) working out the best role for the teachers and getting them on board, (ii) solving the problems with parents who don't understand computers, don't have one at home and so can't help with the homework; (iii) solving the problems with parents and funding agencies who just prefer the idea of having their children taught by people rather than machines; and so on. He doesn't understand the *struggle up* approach the critic needs to employ with technology. So he does not give due weight to the *potential* merits, and how to make sure they are manifested. The book we needed from him, or from someone else, was ***Overcriticized and Underused: Benefitting from Computers in the Classroom***.

The hardware technologists, led by Nicolas Negroponte at MIT, saw the analogies to all this in thinking about computers for schools, and found their way around the problems in order to pull off the brilliant design that is the $150 computer.[34] But we've missed the same point about the software. The trick here is exactly the same as it was in thinking about the heated filament idea for Edison's light bulb: a great idea, but *thousands of steps from the idea to success*.

The same problem arose for developing the content of Skinner's teaching machines, or the content of the programmed texts, or for the basic idea of the hardware: you have to start with really smart ideas and then try some version of them and make really smart improvements of them, not just once or twice or three times but as many hundreds of times as it takes until they do the job or demonstrate that they can't do it. That's how nature does it, too; the advantage nature had was millions of years instead of decades. We're moving too slowly ahead of the tsunamis of dictators or terrorists with weapons of mass destruction, and overly crude reactions to them.

The magic of CAI, or of PTs, lies in the fact that with enough smarts, and enough hard work, and serious, critical, evaluation of every aspect of every revision, and using the smarts to solve every problem the evaluation turns up at every step, the technology *can* do the job. That was clear enough from the teaching machines that became the PTs, and the talking typewriter study sixty years ago, and demonstrated again with the success of the PTs—and

[34] The associated project, his vision, is the OLPC (one laptop per child) project, and it has had plenty of demonstrated success at the village level, in getting the children web-literate; but not yet literate. It is the perfect complement to the Bork project

PSI and HIP—in the medium scale studies.

But to get to the large scale application—to the public schools of a state—you have to have reasonably well-educated funders to provide the development support, people who understand that technologies have to be supported hard and long once they have demonstrated that they work on a small scale—as long as no insuperable obstacles emerge in the mid-scale development. And you have to have good evaluators of the mezzo level efforts, to help with identifying the still-needed improvements and the solutions to the remaining problems.

For you still do have to solve some serious problems. These start with getting the media educated, since it can help with the rest and can seriously handicap everything else if it's *not* supportive; getting continuing funding for development from government, business, philanthropies, or public subscription;[35] getting parental support, getting teacher support, and getting student support. Each of those groups will need to see good articles, good books, good evidence, and good debates at their interest level.

Even with good evaluators on board, it's clear the absence of the basic education of the citizenry about technology makes the road a hard one. Still, if we had even had enlightened funding agencies—private, commercial, or public—we might have been able to do it. But we had such a pervasive dominance of the wrong concept of technology that even good evaluators were missing. The only bright light in the entire scene was the maverick visionaries, from Moore and Skinner and Keller to the PT, PSI, and HIP groups on the educational side, and the OLPC (One Laptop Per Child) group on the hardware side. They created and demonstrated the working model, the 'proof of concept,' and the rest of the country let it die.

Keeping it going was what the 'struggle up' concept means today, just as it meant when our ancestors started to bang rocks around, and finally got them shaped into elegant arrow points, or converted grunts and gestures into basic language. It took them a million years or so, and a ton of smarts; no one should complain if it takes a few years to get good voice-input/output software in Swahili that really makes it possible for computers to teach kids to read and write. But now it's sixty years, the impossible hardware is in place, and all we're doing is writing obituaries: for that situation, the only description that fits is the decline of the best.

Keep in mind that the hardware for doing this for our own children was always in place; Skinner was doing it with books and Moore with computers in the mid-twentieth century. What's happened since then is that the computer visionaries have now given us the hardware for globalization—that is, for doing it in the remote villages of Cambodia, where Negroponte first used the OLPC computer, or the *favelas* of Rio, where the Brazilian government is testing it as this is being written.

There's no need for the non-technologists to stand idly by while the development for the macro scale version goes on, of course; they could have been working on solving the se-

[35] This sounds a bit odd to American audiences, but one of the most important medical research groups in the world is the European group, funded by public donations, that exposed the corruption in FDA drug testing a few years ago.

rious problems with converting the larger context from resistance to support. It was indicated above how one might solve the problem of using the best the teachers can provide along with the best the computers can provide; this can be a real win/win situation, with the teachers moving from the frustration of struggling to apply traditional approaches to mixed-level classes to the great successes of assisting individuals where the computers or programmed texts haven't been successful, and helping authors rewrite the programming to avoid these roadblocks with the next edition. Only in this way can we hope to engage teachers fully in understanding the individualization approach and helping it come to fruition. This first step can help with the other contextual problems to be addressed with teachers, e.g., about the lack of freedom for the teacher to adjust the content, which many of them resent. The solution here is perhaps to stress that there will be plenty of time for teachers to do it their way with the students that get through the required basic content fast. For the slower learners, it is hard to justify deviations from the core curriculum, with plenty of room for individual help getting them through it.

Then there is the problem of students finishing the same curriculum at different times; a little ingenuity from the administrators as well as the teachers is needed there, and a little flexibility. Technology requires some struggling up by those who wish to benefit from it; it does not always provide a free lunch.

What about the parents, who often express their concern at having their children 'taught by machines instead of a caring person'? There's no short road to winning them over; perhaps the best way to do it is to have them test drive the compututor, watch their children use it, and then discuss what they've seen. Of course, if they are already used to their kids playing games with computer opponents, it will be much easier to get them to adjust to Hal as a tutor. But in reality they may already feel threatened by being left out of the game playing, and, unable to forbid their kids from having fun, they want to show they are still powerful enough to control what happens to their kids in school. These are all fixable problems, but they are all fatal problems if we don't work on them and instead just point to them as not being solved by computers.

If we had more space here, it would be worth reviewing the other half of the great disaster in the history of educational technology from the talking typewriter to the present. O. K. Moore had a brilliant double vision, those many years ago; the first half was the computer tutor, the second, his 'autotelic' theory, was that all learning could and should be made into fun. Computer games are the best way to teach many of the most important and most difficult things in the curriculum—scientific method, for example, or the nature of technology. They provide incomparable motivation, because they're fun; and, suitably programmed, can supply incomparably fast feedback to the player's responses, because they are fully individualizable to the player's skill level. Again, the teacher's role is vital for maximizing the educational content of what can be taught in this way. [36]

Another example of that point is educational TV; the research makes clear what fast and enjoyable learning is possible from television, even from the standard shows on the nature,

---

[36] I spelled this out in a long article published 20 years ago, when it was a new idea: "Taking Games Seriously" in ***Educational Research & Perspectives*** , Vol. 14, No. 1, June, 1987, and also on my website if you use Google to find it (the exact address is about to change).

science, and history channels, with viewing assigned as homework, *if* properly orchestrated with advance questions to be answered and discussed with the teacher the next day. Sesame Street doesn't teach much, if anything, as the serious evaluations show, any more than mere viewing of TV shows on forensic science does; but TV *can* teach a great deal. This approach is interactive, although in a different way from the games. But the barriers to it are much the same, for example in terms of parental opposition to using 'what should be seen as entertainment' instead of 'real homework.' Ladders over those barriers have to be invented, built, and maintained, too.

One step away from games, but retaining some of their pleasure and the most important of their technological characteristics are computer simulations. Long used in the military, and to some extent in business schools and military training, they can provide a critical bridge over the gulf between classroom and reality. In training fighter pilots they have saved millions of dollars in crash reduction, and in training civil aviation pilots, millions of dollars in supervised flight time. But they could make larger savings in training young drivers; we just have not been willing to put up the money to handle the software and hardware requirements, so we stay at 40,000 deaths per annum on the US highways and byways, of which a hugely disproportionate fraction is young drivers.

Of course, some of this is the effect of drink, drugs, and testosterone; but some of it is simply due to lack of adequate realistic training, which is, to put it bluntly, simply absent because we can't afford to provide enough one-on-one driving instruction, especially in adverse weather and traffic conditions, for every child in the country to become truly adept before they are turned loose to use the public highways as their learning zone. Now *that* problem can be fixed using supersimulators, i.e., simulators with screens providing a 180º field of view, surround sound, and seats that simulate inertial effects. We know, from the cost of the supersimulators now on the market for game aficionados playing Gran Turismo, the most successful computer game of all time, that the cost of that kind of equipment can be brought down below $2000 per station in quantity production.

Now suppose, worst case, that these machines last 25 months of 50hrs/wk, a schedule that would be easy to fill in terms of the demand, and that maintenance, replacement parts, and power costs for running them cost, worst case, $480/yr.—that comes to about 50 cents/hr/machine. Paying rent on housing these in a low-rent area on public transport, and paying a supervisor for a large room containing ten of them (you could probably get volunteers for this) adds at most $3.60/hr/seat, which gives us a total of around $4/hr. Compare that with the cost of an instructor and a car for each learner at about $40/hr (minimum), the present alternative that provides brief individualized instruction in very limited road conditions, and you're looking at a cheap way to provide the learning time needed to pass a really serious driving license test, which appears, from the sketchy evidence available, to be about 100 hours.

A serious test means the usual inadequate test to save having the change the law, *plus* a test on the supersimulator with lots of problems like skids on rainy roads, defensive driving at night on winding two-lane gravel roads with big trucks going each way at variable speeds, handling punctures, breakdowns, and a car full of yelling teenagers/babies/dogs, etc. Of course, we need a foundation to fund the initial cost of the software, but there are many who would do this.[37] And we would need to test this approach before investing on a large scale, but it's another example of the need to use technology to get over the Rockies, meaning the reality of learning to drive, instead of using personal instruction to learn how to cross the street. Almost any of the large urban school districts in the US, or a federation of smaller ones in the same state, could fund this approach, but we neither do it nor investigate it seriously, despite the annual death toll.

Note that in any discipline where the students need field experience—and evaluation is a leading example of this—the use of supersimulators would take us a quantum jump beyond what we can otherwise afford, with results that may be—we need to find out—nearly as good as field experience. At the moment, in the doctoral program I ran at Western Michigan University, we were lucky enough to have a client who has been well served by advanced students working in the villages of distant lands; but I wish I could have trained them more, before they got to the field. And that field is only one of many that we could simulate well enough so that they would acquire valuable learning from working on the simulator. This is surely the future of professional training, for surgeons and ski instructors, technicians and technologists.

Conclusion

As the reader will have noticed, one sad moral of this story is that the evaluators also failed, and that means those who trained them failed. They, too, are not clear about technology, neither the concept itself nor, consequently, the evaluation of it. Very few of them get any training in technology assessment, which is a long way from program evaluation. When you are evaluating a program for putting computers into education, which means almost any program in education (since good evaluation should consider alternatives to the program under evaluation), it's not enough to be just a program evaluator.[38]

So the argument goes like this. In the mid-twentieth century we had in our hands the key to transforming most of education, and people with the vision to develop it. Some variants of this were developed on the small scale that showed it was not wedded to computers as such, and could be developed in book form, or lecture form, or in electronic media unlimited by distance. But in fact we dropped the key, and this failure was catastrophic in three ways. It was in the first place a glitch in developing a curriculum and pedagogy that would have made our educational system more effective and more enjoyable for those in it; and, make no mistake, the evidence is clear that this means it would have saved many of the millions the present system has served badly or not at all from significant levels of unemployment and poverty (and deaths on the road). But it also happened that the greatest impact of this failure was on the slice of education that should have provided the educated personnel and inspiration needed to maintain and improve the technological foundation of our socie-

---

[37] I have worked with the group of foundations interested in this at a meeting funded by the AAA Road Safety Foundation to address the reduction of road deaths.

[38] For an outstanding example of good evaluation of a technology program, see Denis Gray's paper in ***New Directions in Evaluation*** (Summer, 2008), ***Reforming the Evaluation of Research***, C. Coryn and M. Scriven, eds.; Jossey-Bass, 2008.

ty. The vast cost of this failure to all of us, including those who succeeded in the outmoded educational system we perpetuated, was loosely predictable and is now becoming apparent. Third, it also happened that the path we chose not to develop was the best, and perhaps the only, hope for any kind of universal education in at least a third of the rest of the world. The consequences of this for the failure to provide the underpinnings of democracy there we are also now beginning to see in full bloom. There is also the need to improve access to professional-level work in evaluation, an important tool for developing leverage for upgrading the globalization of education: here we are a long way ahead of our colleagues in educational research, but could still do more.

Our failure was due at its heart to three flaws in our society. The first was the complete failure to understand technology, a failure massively encouraged by the science lobby and not corrected by the educational experts, who did no better because they didn't bother to think hard about technology, or didn't like their chances in fighting the science lobby. The second is the failure of our society to fully understand that education, if done well, is the fundamental survival mechanism of contemporary society, not just a budget item to be cut along with everything else when times get hard. It is the one thing that in the long run makes everything worthwhile possible, by making all people able to play a meaningful part in their own governance. The third was the failure to insist on getting good 'evaluation by results' in education, to insist on using technology to get more of these results and to broadcast them globally, and to stick by the lessons implicit in those results through the long process of bringing people out of old prejudices or ignorance to find and face the true in the new.

All of us share the responsibility for these failures, and the best to be hoped for from this essay is that it may encourage reconsideration of the extent to which they have cost us all dearly, and cost our grandchildren, and cost our fellow citizens of the globalized village.

REFERENCES & ACKNOWLEDGMENTS

I give no general supporting references because I am here mainly reflecting on common (i.e., Googlable) knowledge, or personal experience, or diverging radically from the available published interpretations. One of my first assignments when I joined the faculty at the University of Minnesota in 1952 was the course in history of science and technology, and I have continued to work in that field ever since. I learnt from many sources, beginning with Conant and the Cornell historians, the source of much of Kuhn's relativist variation on their theme, and then from my colleagues in the NSF-founded Department of the History and Philosophy of Science at Indiana University. I have however been mainly influenced in my reading of the history of technology by the training in conceptual analysis that I received from those influenced by Wittgenstein, more so than most of my colleagues in the philosophy of technology, except perhaps Norwood Russell Hanson, who was academically, although not personally, more interested in mainstream science than technology. And I have learnt a great deal from working as an evaluation consultant in the technology start-up funding arena.

Special thanks, too, to both the editors of this volume and the anonymous reviewer from Australia for helpful criticisms and suggestions that vastly improved the original version of this paper.

Patterns of Discovery: An Inquiry Into the Conceptual Foundations of Science. NR Hanson (1958), Cambridge University Press.