MILITARY MEDICINE, 178, 10:37, 2013

Cost Considerations in Using Simulations for Medical Training

J. D. Fletcher, PhD; Alexander P. Wind, MS†*

ABSTRACT This article reviews simulation used for medical training, techniques for assessing simulation-based training, and cost analyses that can be included in such assessments. Simulation in medical training appears to take four general forms: human actors who are taught to simulate illnesses and ailments in standardized ways; virtual patients who are generally presented via computer-controlled, multimedia displays; full-body manikins that simulate patients using electronic sensors, responders, and controls; and part-task anatomical simulations of various body parts and systems. Techniques for assessing costs include benefit–cost analysis, return on investment, and cost-effectiveness analysis. Techniques for assessing the effectiveness of simulation-based medical training include the use of transfer effectiveness ratios and incremental transfer effectiveness ratios to measure transfer of knowledge and skill provided by simulation to the performance of medical procedures. Assessment of costs and simulation effectiveness can be combined with measures of transfer using techniques such as isoperformance analysis to identify ways of minimizing costs without reducing performance effectiveness or maximizing performance without increasing costs. In sum, economic analysis must be considered in training assessments if training budgets are to compete successfully with other requirements for funding.

INTRODUCTION

The advantages of using simulation in training may be summarized as follows:

—Safety: Simulated lives and health can be jeopardized to any extent required for learning.
—Economy: Simulated materiel, equipment, and other resources—physical or fiduciary—can be used, misused, and expended as needed.
—Visibility: Simulation can provide visibility in at least two ways. It can (1) make the invisible visible and (2) control the visibility of details allowing the learner to discern the forest from the trees or the trees from the forest as needed.
—Time control: Simulated time can be sped up, slowed down, or stopped. It can also be completely reversed, allowing learners to replicate specific problems, events, or operational environments as often as needed.

These advantages seem to be applicable in medical training and education as elsewhere. Overall, simulation can provide massive amounts of practice with feedback, exposing individuals or teams to realistic situations that in real-world settings would range from the impracticable to the unthinkable.

All these advantages are relevant and interrelated. This article focuses on the economic value of simulation. It suggests ways to assess the use of simulation in medical education and training through objective economic and cost analyses. The article begins with brief reviews of simulation in medical education and training, and economic and cost analyses. These reviews are followed by a discussion of ways in which measures of simulation training effectiveness can be combined with cost analysis to yield assessments of costs, cost-effectiveness, and return on investment in medical training and Education.

SIMULATION IN MEDICAL TRAINING AND EDUCATION

As abstracted from reviews and comments (e.g., Bradley[1] and Rosen[2]), at least 4 forms of simulation appear to be used in medical training.

Standardized Patients

Applying formal procedures developed in 1964 by Barrows and Abrahamson[3] and continued into the present,[4,5] actors, real patients, or lay people can be trained as “standardized patients” to participate in role-playing exercises for assessing and improving a leaner’s ability to carry out medical procedures, such as taking medical histories, performing physical examinations, ordering tests, providing counsel, and prescribing treatment. These patients can be available when and where they are needed, trained to respond consistently to examination questions, and used when training with a real patient would be inappropriate, as in counseling cancer patients. They are, however, expensive to recruit and train and cannot present cues normally provided by physiological examinations.

Examples:

—Gerner et al[6] found that consultation quality ratings by standardized patients after their visits with 67 general practitioners predicted later ratings by parents concerning improvements in the weight control behavior of their children. 95% of the general practitioners reported that they found training with standardized patients to be useful.

*Science and Technology Division, Institute for Defense Analyses, 4850 Mark Center Dr., Alexandria, VA 22311.

†University at Albany, State University of New York, 1400 Washington Ave., Albany, NY 12222.

The findings and opinions expressed here do not necessarily reflect the positions or policies of the Office of the Secretary of Defense.

doi: 10.7205/MILMED-D-13-00258

—A study by Betcher[7] found that the use of standardized patients (graduate students in theater) in role-playing consultations, followed by debriefing, was effective in improving the communication skills and confidence of nurses and other caregivers by 5% to 37% in advising end-of-life patients and their families.

—Safdieh et al[8] compared the long-term effects on quality of neurologic examinations performed by 58 medical students who were trained using standardized patients to those performed by 129 students who were trained without this experience. Two years after this training, the authors found that a statistically significant advantage in performing these examinations favored students who were trained using the standardized patients.

Virtual (Computer-based) Patients

Interactive software simulations of patients have been used in standard simulation exercises[9,10] and in gaming simulations[11] for training and assessing medical skills. These simulations are gradually taking the place of standardized patients, although the absence of strong artificial intelligence, which would allow full mixed-initiative dialogue, limits their applicability. However, growth in the use of virtual patients is likely to continue because of their ability to scale inexpensively to large numbers of physically dispersed learners, adapt quickly to prior knowledge and other individual characteristics of learners, and be available anytime and anywhere via the global information infrastructure.

Examples:

—Steadman et al[12] compared learning in a week-long acute care course by 31 fourth-year medical students who were randomly assigned to a group receiving a virtual patient with labored breathing versus a patient receiving interactive problem-based learning without simulation. The simulation group performed significantly better (71%—a 24-point gain over the pretest) on the final assessment than the problem-based learning group (51%—a 7-point gain over the pretest).

—Ten Eyck et al[13] used a crossover design involving 90 students to compare virtual patient simulation with group discussion in emergency medical instruction. The learners received one set of topics using one instructional treatment and then switched mid-rotation to receive the other set of topics using the other treatment. Material presented in simulation format produced significantly higher scores than material presented using group discussion methods.

—Botezatu et al[14] compared learning by 49 students studying hematology and cardiology topics using virtual patients with learning by students receiving more conventional instruction (lecture and small group discussion). They assessed learning immediately after instruction and its retention after 4 months. They found effect sizes ranging from 0.5 to 0.8 in favor of the virtual patients.

Electronic Patients

These types of patients are generally whole-body manikins that physically simulate patients, although some use of helmet-mounted virtual reality capability has also been used.[15] Military and civilian simulation developers have collaborated to produce and assess electronic patients. An early effort by David Gaba (Stanford University) and CAE-Link developed a full-scale manikin system called Anesthesia Crisis Resource Management. It was designed for Air Force and National Aeronautics and Space Administration Crew Resource Management training.[16]

Examples:

—Alinier et al[17] compared the performance of intensive-care nursing students who were assigned at random to two groups. One group followed the usual course of training, and the other group used a "universal patient simulator." The researchers found significantly greater performance improvement (twice the percentage increase) for the simulation-trained group than for the group receiving the usual course of training.

—Radhakrishnan et al[18] used electronically controlled manikins to compare the performance of manikin-trained and conventionally trained nursing students. They found that the manikin-trained students significantly outperformed the other students in patient safety and in assessing vital signs.

—Cendan and Johnson[19] used a randomized, repeated-measures design to train 40 second-year medical students in treating neurogenic, hemorrhagic, septic, and cardiac shock. This study compared two instructional approaches such as web-based training text with a culminating simulation exercise and a manikin-based exercise with instructors who provided management and evaluation in response to student questions. All students were exposed to both approaches, with half completing the web-based exercises first and half completing the manikin-based exercises first. Learning from web-based and manikin-based instruction was similar; however, overall learning was greater when the web-based simulation was presented first.

Part-Task Trainers

Anatomical models of body parts are used to provide training. These "part-task" simulations are becoming more advanced to keep pace with medical treatments and technology. They are used in instruction that ranges from minimally invasive laparoscopy[20] to major cardiologic surgery[21] and delicate ophthalmological procedures.[22] They have been enhanced considerably through the development and inclusion of haptic systems and interfaces.[23]

Examples:

—Barsuk et al[24] found significantly fewer needle passes, arterial punctures, and catheter adjustments and higher overall success rates among 76 simulator-trained residents using a haptic task trainer than among 27 residents trained without the simulation.

—Holzinger et al[25] compared learning by 96 medical students who were randomly divided into 3 groups: a conventional text-based lesson group, a group learning from a blood dynamics simulator alone, and a group learning from the simulator blended with additional human instructor support. They found no difference between the first two conditions but significantly more learning in the third.

ECONOMIC ANALYSIS: THE VALUE OF A POUND OF SIMULATION

Administrative decision making largely consists of allocating resources among competing alternatives.[26] Such decision making is not only a matter of adopting enhancements, but also includes determining what must be given up to do that. To some degree, this process is described by the "rational theory of choice," which balances costs against factors that contribute to the achievement of a specific goal.[27–29] This approach can be overdone by neglecting issues that are hard to measure (e.g., attitudes, culture, and trust), but it plays an essential role for cases in which benefits and costs can be clearly identified, such as business profit and loss, combat success or failure, and the achievement of objectives through instruction.

The objective of economic analysis is to inform decisions by assessing alternative courses of action and/or inaction. It does so by estimating the amount and probability of expected returns from each alternative and balancing these returns against projected costs, consequences, and constraints.[30]

Economic analysts look for value to be gained and resources to be sacrificed for each alternative identified. Such deliberation assumes that the analyst or the decision maker has assembled and considered a comprehensive list of available alternatives. Introduction of an additional alternative can dramatically alter the decision space.

Economic analyses remain as subject to controversy as any other analyses. Underlying assumptions, inclusion and exclusion of data elements, proper data collection, sampling procedures, criterion levels, and similar issues contribute to controversy. An economic analysis can never be assuredly correct, but it can and should be explicit and should allow decision makers to determine how well and to what extent it informs their decisions.

The single most accessible and readily commensurable criterion for choosing among alternatives remains cost measured in fungible monetary units (e.g., dollars). In these cases, decisions key on units returned for units invested. Even with these data, the decisions do not make themselves. Other factors poorly suited to economic analysis also come into play. The effectiveness of military and medical training outweighs cost when human lives are in risk. Cost considerations—however necessary, objective, and well-conceived—should not be the sole concern in informing decisions.

Assessing Costs

Assessment of costs invested is a central factor in economic analysis. Costs may be categorized as one of the following: research and development, initial investment, operations and maintenance, and salvage and disposal.[30] Salvage and disposal costs are omitted from many analyses because they are one-time only and difficult (usually impossible) to estimate accurately. Many research and development and initial investment costs are also one-time only, but they may be known and may be a matter of interest for some alternatives. When an alternative is being considered as a replacement for an existing program, both research and development costs and initial investment costs for the replacement can be included even though they are not included for the program in place. In these cases, analysts may decide that the costs for the current program are "sunk" (i.e., beyond recovery no matter what happens). These sunk costs do not factor into the decision.

A perennial and debilitating problem for cost analysis in education and training is the absence of generally accepted, standardized cost models. Such models would present unambiguously specified and well-defined cost elements that clearly identify what they do and do not include. Without these specifications, decision makers, among others, do not know clearly what the cost analysis and the cost analysts are telling them.

A variety of commentators have provided a basis for cost models to be used in instruction. For instance, Levin and McEwan[31] suggested five classes of elements, or "ingredients," to be considered in a cost model: personnel, facilities, equipment and materials, other program inputs, and client inputs.

Personnel costs include all the human resources needed by the approach. Levin and McEwan recommend that personnel be classified according to their roles (instructional, administration, clerical, and so forth), qualifications (training, experience, specialized skill), and time commitments (full time, part time). Facilities costs include all resources required to provide physical space for the approach. Equipment and materials include furnishings, instructional equipment, and supplies. Other inputs in this scheme include components that do not fit elsewhere (e.g., instructor training and insurance costs).

Other costs are especially relevant in military and industrial training, where student pay and allowances are funded by the same organization that provides the instruction, thereby increasing interest in the speed with which students reach objective thresholds of competency. Much of the rationale for applying technology in industrial and military training is keyed to its capacities for tutorial individualization, which allows the adjustments for prior learning and self-pacing to qualify

	Personnel	Facilities	Equipment	Materials	Indirect Costs
Analysis					
Design					
Development					
Implementation					
Evaluation					

FIGURE 1. Cost model framework for instruction.

students more quickly for duty or allows students to maximize their competencies—be all they can be—while holding instruction time constant.[32]

Kearsley[33] developed a model much like Levin and McEwan's but with an added dimension for the components, or categories, of instruction system development: analysis, design, development, implementation, and evaluation. These components can be combined with the typical cost categories of personnel, facilities, and equipment and materials. Integrating these two categories yields the cost framework shown in Figure 1, which presents an outline, not a fully developed cost model. Explicit discussion of what is included in, and/or excluded from, each cell of this framework will help analysts know what they are talking about and will help decision makers determine the extent to which an analysis can be applied to inform their decisions. It is rare, if not impossible, for a cost analysis in any area, including instruction, to be entirely correct. As discussed earlier, every such analysis requires assumptions and extrapolations, but it can and should be explicit. The framework shown in Figure 1 may contribute to this end.

Benefit–Cost Analysis

A benefit–cost analysis is used to determine whether the benefits returned by a candidate course of action outweigh the costs of investing in it. The calculation of a benefit-to-cost ratio is straightforward as described by Fitzpatrick et al[34] Phillips,[35] and McDavid and Ingleson,[36] among others. It reduces all costs of an action to a single unit. It does the same for all benefits and then calculates the ratio of benefits to costs.

We can calculate a benefits-to-cost ratio using whatever metrics we choose, but the terms for input and output must be commensurable (i.e., both must be measured using the same units). Monetary units tend to be those most readily translated from whatever investment resources are required and whatever returns are produced. For that reason, these ratios are usually expressed in terms of dollars, pounds, euros, or whatever monetary unit communicates most easily and usefully to likely decision makers.

A benefit-to-cost ratio is calculated as follows (Phillips[35]; McDavid and Ingleson[36]):

$$\frac{\text{Value of the result}}{\text{Cost of the investment}}$$

It tells us how many units of value we get for every unit of cost.

For instance, Thompson[37] reported that in 1667 public health officials in London found that expenditures to combat the plague would yield a benefit to cost ratio of 84:1.

Return on Investment

Return on investment is closely related to benefit-to-cost ratios. It is also a ratio, and calculating it is as straightforward as its name suggests. It is calculated as follows (Phillips[35]; McDavid and Ingleson[36]):

$$\frac{\text{Value of the result} - \text{Cost of the investment}}{\text{Cost of the investment}}$$

Return on investment must be calculated for some period of time, such as a year. As with monetary units, the length of time should be determined by analysts in consultation with decision makers who are likely to use the results of the analysis.

Example:

—Fletcher and Chatham[38] studied returns from investing in several training innovations. They found ratios of 2.49 for the "Top Gun" investment in training Navy combat pilots, 3.37 for using technology-based, in-transit training to sustain and enhance the bombing skills of pilots, and 2.50 if technology-based training were used for 40% of Department of Defense specialized skill training.

Benefit–cost and return-on-investment analyses require value and cost to be commensurable. Of the two, return-on-investment analysis may be preferred because it indicates how many units of net benefits are returned, after investment costs have been subtracted, for each unit invested. Of course, spikes, dips, and diminishing returns have to be considered with differently timed units of investment, so averaging and curve smoothing may be required.

Return-on-investment analysis may be helpful for an ancillary reason. It treats costs for education and training explicitly as investments, not as infrastructure expenses. Treating these costs as infrastructure expenses is often their fate in training venues, including those of Department of Defense, where training is bundled with transit, hospitalization, and stockade costs.

Cost-effectiveness analysis

When commensurability is difficult, cost-effectiveness analysis can be used.[31,35,36] Costs of investment can usually be expressed in monetary units, but the full return—the benefits—of instruction may not be amenable to monetary units. Cost-effectiveness analysis allows effectiveness (e.g., information retention, job knowledge and motivation of workers, supervisor ratings, and productivity) to be measured in its own units. In instruction, it accommodates a more comprehensive range-of-objective outcome than analyses requiring commensurability.

Cost-effectiveness is calculated as a direct ratio of cost to benefits or benefits to cost. In determining cost-effectiveness, the usual practice is to hold either costs or effectiveness constant across all alternatives being considered and observe variations in costs or effectiveness. Sometimes, either costs or effectiveness is simply assumed to be constant across the alternatives. One could argue that cost is implicitly assumed to be constant by its absence from many instructional evaluations. The assumption may be reasonable, but analysts should present data or information to validate it so that decision makers can decide for themselves if it is warranted.

The good news is that cost-effectiveness does not require commensurability. The bad news is that it is a relative term. Relevant decision alternatives must be specified in assessing it. The addition of an alternative for achieving the objective(s) after a cost-effectiveness analysis is done can change its conclusions and recommendations entirely. Despite common usage, we cannot properly say that an investment, by itself, is or is not cost-effective; however, no harm is done in calculating a cost-effectiveness ratio for it.

Example:

—Fletcher et al[39] combined experimental data reported by Jamison et al[40] and Levin et al[41] with their own empirical findings to assess the costs of raising student scores on a standard test of mathematics comprehension by one standard deviation. They compared these costs for professional tutoring, peer tutoring, reducing class size, increasing instructional time, and using computer-based instruction. They found that the most cost-effective approaches among all these alternatives were peer tutoring and computer-based instruction.

Cost-effectiveness analyses have been used in health care since the 1960s to determine the relative value of specific interventions, such as a medication, surgical procedure, or counseling techniques.[42]

Example:

—Tsai et al[43] calculated cost-effectiveness ratios that compared hospital-based home care for patients who had mental illness with care based on traditional, outpatient therapy. They measured effectiveness in terms of disease maintenance behavior, psychotic symptoms, social function, and service satisfaction. Overall cost was the sum of costs for all direct mental health services. They found cost per unit of effectiveness to be \$4.3 for home care and \$13.5 for outpatient therapy.

One form of cost-effectiveness analysis is cost-utility analysis, where the return is assessed in terms of utility or value received by the beneficiaries of the investment. Cost-utility analysis is frequently recommended and promoted but rarely used in sectors other than health services, where decision makers often assess different quality-of-life alternatives for their patients.[44,45] They must balance quality of life against additional years of life to help patients review the net benefit or utility provided by different treatments.

Assessment of Simulation

Decision making concerning potential improvements in training raises two basic questions. Compared to current practice, does it produce threshold levels of human performance capabilities at less cost, or does it increase human performance capabilities while holding costs constant? Both costs and effectiveness must be considered if assessments of training simulation are to inform decision making in a responsible manner.[30,46]

Transfer effectiveness ratios

A key issue is the extent to which capabilities produced through simulation-based training transfer to "real-world" tasks. More specifically, does the human performance produced by simulation-based training either reduce costs without diminishing performance or improve performance without increasing cost? One approach to this issue is the use of transfer effectiveness ratios (TERs). TERs were developed by Roscoe and Williges[47] for aircraft pilot training, but they apply to simulation-based training in general. A TER can be defined as follows:

$$\mathrm{TER} = \frac{T_c - T_x}{X},$$

where TER is the transfer effectiveness ratio; T_c is the time or trials required for a control/baseline group to reach criterion performance; T_x is the time or trials required for an experimental group to reach criterion performance after X time or trials using simulation (or any other instructional approach of interest); and X is the time or trials spent by the experimental group using the simulation.

Roughly, the TER indicates how many trials or units of time are needed to achieve criterion performance in the objective experience (e.g., flying an aircraft, repairing a radar repeater, and performing a medical procedure) are saved for every unit of simulation training invested.

Example:

—Taylor et al[48] used TERs to compare times required to reach criterion performance in using specific aviation instruments with and without a Personal Computer Aviation Training Device (PCATD). One group was trained only during flight in the aircraft. A second group was trained first with the PCATD and later in the aircraft. Criterion performance was measured during flight. Taylor et al found that the PCATD group required about 4 hours less of in-flight training, suggesting a transfer effectiveness ratio of 0.15—or a savings of 1.5 flight hours for each 10 hours of PCATD time.

These findings suggest that the requisite levels of performance can be attained at lower cost using simulation—if the simulation costs less to operate than an airplane. If it does, then "the larger the TER, the better" is good news for the simulation.

Example:

—Orlansky et al[49] compared the costs of flying military aircraft with the cost of operating ("fly") simulators. They found that the cost of operating a flight simulator was about one-tenth the cost of operating military aircraft, so the use of a flight simulator was generally cost-effective if the TER for the simulator exceeded 0.10.

This finding is useful and significant. However, a few caveats are in order.

First, as Povenmire and Roscoe[50] pointed out, not all simulation training hours are equal. Early trials or hours in a simulation may save more trials or time than later ones. A TER is likely to decrease monotonically and approach zero for large values of simulation training. This consideration leads to learning-curve differences between TERs and incremental TERs, or ITERs, with the inevitable diminishing returns captured best by the latter. An ITER can be defined as follows:

$$\text{ITER} = \frac{T_{x-\Delta x} - T_x}{\Delta X},$$

where ITER is the incremental transfer effectiveness ratio; $T_{x-\Delta x}$ is the time or trials required to reach criterion performance with access to simulation after completing x–Δx units of time or trials; X is the time or trials spent by the experimental group using the simulation; T_x is the time or trials required to reach criterion performance, with access to simulation; and ΔX are incremental units of time or trials after starting at unit X.

Roughly, the ITER indicates the amount of transfer produced by successively greater increments of time or trials in the simulation. As Morrison and Holding[51] pointed out, total time or trials to criterion begins to decrease as the use of effective simulation increases, but, sooner or later, they begin to increase. At some point, training total time or trials to criterion with simulation will exceed those without it and produce negative TERs.

Example:

—Taylor et al[52] used ITERs to compare the number of trials to specific completion standards, time to complete a flight lesson, and time to a successful evaluation flight with and without a PCATD. One group trained only during flight in the aircraft, and three other groups trained first with the PCATD and later in the aircraft. Criterion performance was measured during flight. The number of trials to reach criterion was less for all three PCATD groups than for the aircraft-only group. The three experimental groups trained with the PCATD for 5, 10, and 15 hours, respectively. The 10-hour PCATD group required the fewest number of trials to reach criterion for five of the eight criterion tasks, the 5-hour PCATD group required the fewest number of trials for two of the criterion tasks, and the 15-hour PCATD group required the fewest number of trials on only one criterion task. Average ITERs were 0.662, 0.202, and 0.148, respectively, for the 5-hour, 10-hour, and 15-hour PCATD groups, indicating the diminishing returns from time in simulation training accounted for by ITERs.

Second, transfer effectiveness is tied to the specific skill, knowledge, or performance levels—the training objectives—being sought. This issue was illustrated in a study by Holman[53] involving a CH-47 helicopter simulator. Holman found that if the knowledge and skills of interest were simply overall ability to fly the helicopter, the TER was 0.72. However, he also found that the 24 TERs for the specific skills he examined ranged from 2.8 to 0.0.

The TER that is relevant depends, as in all assessment, on the decision it is intended to inform, which includes the type of transfer sought. Holman required straightforward, "near" transfer, where many elements exercised by the simulator are similar, if not identical, to those required by the objective task performance. Near transfer echoes long-standing prescriptions for including identical elements that are shared by the learning (e.g., simulation) and the eventual task environments.[54,55] Other applications may require "far" transfer from simulation to the objective task, where fewer elements are common to both simulation and task performance and higher level thought and analysis is required by the performer for transfer to occur.[56]

Similarly, transfer may key on automated responses learned in simulation to the objective environment in a straightforward "low-road" fashion. Alternatively, it may require less focus on automatic responses and greater abstraction of simulation performance to conceptual levels that transfer indirectly but broadly to many objective environments (i.e., "high-road" transfer, which requires purposeful attention in the learning environment to the development of learners' transfer abilities[57]).

Given the physical and anatomical differences of human beings, both low- and high-road transfer seem particularly important in the development and assessment of medical training. These forms of transfer have received some attention in the development and assessment of simulations for medical training, but more may be in order.

Third, the operating costs of objective, targeted tasks differ markedly and can produce quite different tradeoffs in assessing the cost-effectiveness of simulation-based training. For instance, Povenmire and Roscoe[50] considered flight simulation training for Piper Cherokee pilots, where the cost ratio of simulation to targeted performance was 0.73, thereby requiring a much higher TER for cost-effectiveness than Orlansky et al[49] found for high-performance military aircraft.

Cost/effectiveness versus effectiveness/cost

TERs primarily concern ways to minimize costs while holding effectiveness constant. In this regard, the ratio between the fastest and slowest learners in typical classrooms appears to be at least 4:1.[32] Learner ability remains a factor, but this ratio is most directly linked to prior knowledge. The variety and extent of prior knowledge increases with the age and experience of the learner, thereby making adjustments for it increasingly important for adult learners. Individualizing instruction by taking into account the knowledge and skill that each learner brings to training has been found to reduce time or trials to reach criterion levels of performance. Costs for specialized technical training in areas such as medicine might be reduced by as much as one-fourth if the capabilities currently available through computer technology were implemented to take advantage of these differences.[38]

On the other hand, maximizing effectiveness while holding costs constant may be more appropriate for military training. Personnel commands, which prepare orders to pass course graduates on to their next duty station, have found it prohibitively difficult to deal with individuals leaving training at arbitrary times. Fast learners who finish early are often detailed to necessary but undesirable duties and thereby have few incentives to save resources by shortening their time in training.

It appears to be more feasible and beneficial for military organizations to provide training that allows each learner to "be all they can be," while holding graduation dates for all students constant. For instance, learners who have experience with a topic exercise in simulation might be presented a more difficult exercise on that topic to enhance their knowledge or skill while holding simulation time constant. This procedure could accommodate the needs of military personnel systems to synchronize the preparation of post-training orders and, through various personnel actions, provide incentives for learners to take full advantage of opportunities to train beyond threshold levels of performance—training that could best be made available using simulation.[32,38]

Cost savings under procedures to maximize performance while holding time constant have been shown to be considerable.[38,58] Unfortunately, most of these savings are realized in duty commands and not in the training commands that must bear the costs of developing and providing the extra training for fast learners. These costs can be minimized using simulation, but, at present, local training commands have limited incentives to implement such procedures.

Further, return on investment appears to be relatively insensitive to development costs at military training scales.[58] The Services could invest much more in the development of high-quality training and still receive strong monetary return on investment. The return to operational effectiveness for this investment is also likely to be substantial, but it is far more difficult to assess.

Isoperformance

TERs cover transfer issues, but we would like to cover costs along with transfer effectiveness in a single omnibus analysis so that allocations of training time or trials between, for example, simulation and "hands-on" exercises produce targeted levels of performance at minimal cost. Isoperformance provides one approach for solving this problem. The basic idea is to devise a function, usually depicted as an isoperformance curve, showing every point where different combinations of training inputs produce equivalent performance outputs.[51,59,60] The solution, then, is to find the point on the curve where costs are minimized.

Isoperformance relates two or more training inputs to a training outcome held at some prescribed value or level. It is generally assumed that each input by itself could produce the desired-level outcome; however, some inputs may provide unique contributions to the outcome, necessitating their inclusion at least to some degree. Isoperformance identifies all combinations of the inputs needed to produce the objective performance.

Bickley[61] pointed out that cost considerations in simulation-based training require at least 2 component functions. First, a function is needed to relate simulation trials or time to their costs. Second, a function is needed to relate costs to trials or time in simulation to performance on the "real" task or job.

The first consideration can usually be treated with a simple linear function to account for time or trials in simulation. The second consideration is more complicated. It is called an isoperformance curve because it trades off simulation time or trials with real task experience while holding performance on that task at some threshold level. It requires the analyst to specify a criterion level of performance and a level of confidence for achieving it. Given these considerations, the factors that determine criterion performance can be traded off against one another, as shown in Figure 2.

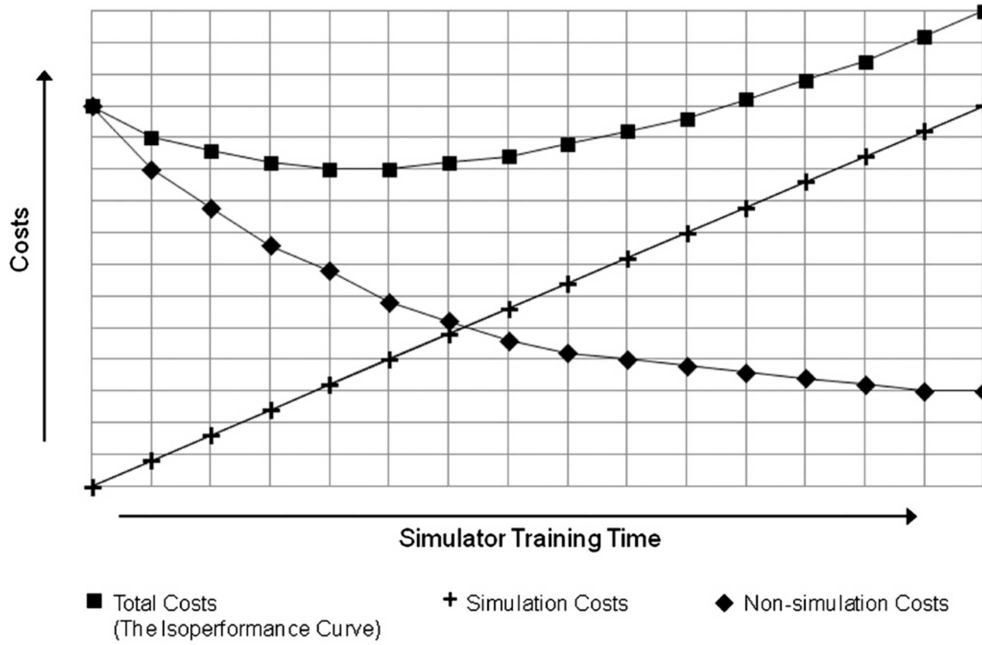

FIGURE 2. Notional isoperformance curve drawn as a function of simulator and actual equipment costs.

Performance—the output of the training—is expected to be the same everywhere on the total cost curve (the upper curve in Figure 2). Total costs initially decrease as simulation time or trials are substituted for those with the (presumably more expensive) real-world objective task or job. Costs then begin to increase as more and more simulation training is allocated and substituted in. Costs for nonsimulation training—initially, the middle curve in Figure 2—start in the same place as total costs but then decrease monotonically as more and more simulation training is substituted in. Notably, these costs rarely reach zero because sooner or later training will have to include time or trials in performing the objective real-world task or job. Costs for simulation—initially, the bottom curve in Figure 2—start at zero and rise monotonically with its increasing use.

Bickley recommended the following formulation for an isoperformance curve, which appears as the top curve in Figure 2:

$$Y = a\mathrm{e}^{-bx} + c$$

where Y is time or trials in the real task required to reach criterion performance; x is the time or trials in simulation; and a, b, and c are the parameters of the model.

Given a data set with reasonable variability in matching simulation time or trials to task proficiencies, the values for a, b, and c in this model can be calculated, and an appropriate isoperformance curve can be developed. Algorithms for doing so are available, as Bickely[61] and de Weck and Jones[60] point out.

The cost-effective solution under this formulation is given by the minimum on the upper, total cost curve. It can then be used to allocate training time or trials between simulation and the real equipment. In effect, it holds performance (or effectiveness) constant and suggests an allocation of inputs that minimizes costs. Carter and Trollip[62] illustrated the other side of the coin. They used a mathematically equivalent approach to devise an optimal strategy for maximizing performance (or effectiveness again) given fixed costs.

The problem of collecting appropriate transfer data to use in TER or isoperformance analyses remains for some applications. Many of these analyses trade off simulation for training (e.g., aircraft piloting and tank gunnery) that otherwise would be exorbitantly expensive. Collecting adequate data to show all combinations of training inputs (e.g., simulation and aircraft piloting) that produce equivalent performance outputs can easily swamp a training developer's budget.

Morrison and Holding[51] suggested that a solution to this problem would be to use limited but valid empirical data accompanied by expert judgment to double-check findings and fill in gaps. They suggest pilot "dosage" experiments with no simulation training, a great deal of simulation training, and two to three different allocations of simulation training in between. Findings from such experiments could then be reviewed and supplemented by expert judgment to produce an approximate learning curve sufficient for either TER or isoperformance analysis. If times or trials to criterion in simulation are a matter of hours or days, if the training is for a critical task or job, and/or if the tasks to be learned are inexpensive relative to piloting military aircraft (as may be the case for many medical procedures), this approach seems reasonable and, in fact, prudent.

Morrison and Holding's[51] application of isoperformance analysis concentrated on gunnery training. The main idea was to use simulation to save training ammunition. Other examples are available. For instance, Bickley[61] focused on simulator versus flight time in the Army's AH-1 helicopter. Jones and Kennedy[59] discussed trading off personnel aptitude against training time. They also provide an appendix that shows step by step how to create an isoperformance curve. de Weck and Jones[60] provide examples from spacecraft design and professional sports.

Isoperformance analysis can be applied to any trade-off issue, including the use of simulation in medical training. Basically, isoperformance curves are just cost curves. It may be time to invest more seriously in this approach.

SUMMARY AND DISCUSSION

Discussion in this article has ranged from generally applicable techniques (economic analysis) to those techniques specifically focused on simulation-based instruction (TERs and isoperformance). Issues of benefit cost, net benefit cost, and cost-effectiveness seem applicable in a straightforward fashion to any sort of medical training and education. However, commensurability is a problem: How are we to capture fully in monetary terms the value of a patient's life, quality of life, and overall health? It is solved to an appreciable degree by cost-effectiveness analysis, provided that we identify a comprehensive set of realistic alternatives.

In contrast to cost-effectiveness, return on investment does not require the identification of all likely alternatives. Different returns from different investments in education and training can be compared later as they arise. However, return on investment focuses on investment costs, which may be unknown or sunk compared to existing alternatives. The result is that the research and development costs and initial investment costs of a new approach may need to be included. The new approach may then be at a disadvantage when considered and compared with return from existing approaches, where such costs are unknown, sunk, and omitted from the analysis.

Applications of TERs and isoperformance to provide economic analyses for simulation used in medical training and education seem both feasible and worthwhile if our analyses are to treat costs in training seriously. Adequate policies and procedures for the cost and effectiveness of training programs might be developed without the expenditure of time, effort, and cost required for optimization, but they will require generally accepted cost models with well-defined cost elements, including those associated with simulation for medical training. These approaches may earn their keep by advancing the field beyond guesswork and/or administrative fiat in the competitive allocation of increasingly scarce resources to medical education and training.

ACKNOWLEDGMENT

Funding for this article was provided by the Office of the Deputy Assistant Secretary of Defense (Readiness), Training Readiness, and Strategy Directorate.

REFERENCES

1. Bradley P: The history of simulation in medical education and possible future directions. Med Educ 2006; 40(3): 254–62.
2. Rosen KR: The history of medical simulation. J Crit Care 2008; 23(2): 157–66.
3. Barrows HS, Abrahamson S: The programmed patient: a technique for appraising student performance in clinical neurology. J Med Educ 1964; 39: 802–5.
4. Collins JP, Harden RM: The Use of Real Patients, Simulated Patients, and Simulators in Clinical Examinations (AMEE Medical Education Guide, No. 13). Dundee, UK, AMEE, 2004.
5. Wilson L, Rockstraw L (editors): Human Simulation for Nursing and Health Professions. New York, Springer, 2012.
6. Gerner B, Sanci L, Cahill H, et al: Using simulated patients to develop doctors' skills in facilitating behaviour change: addressing childhood obesity. Med Educ 2010; 44(7): 706–15.
7. Betcher DK: Elephant in the room project: improving caring efficiency through effective and compassionate communication with palliative care patients. Medsurg Nurs 2010; 19(2): 101–5.
8. Safdieh JE, Lin AL, Aizer J, et al: Standardized patient outcomes trial (SPOT) in neurology. Med Educ Online 2011; 16(1): 1–6.
9. Cook DA, Triola MM: Virtual patients: a critical literature review and proposed next steps. Med Educ 2009; 43(4): 303–11.
10. Cendan JC, Lok B: The use of virtual patients in medical school curricula. Adv Physiol Educ 2012; 36(1): 48–53.
11. Cannon-Bowers JA, Bowers C, Procci K: Using video games as educational tools in healthcare. In: Computer Games and Instruction, pp 44–72. Edited by Tobias S, Fletcher JD. Charlotte, NC, Information Age Publishing, 2011.
12. Steadman RH, Coates WC, Huang YM, et al: Simulation-based training is superior to problem-based learning for the acquisition of critical assessment and management skills. Crit Care Med 2006; 34(1): 151–7.
13. Ten Eyck RP, Tews M, Ballester JM: Improved medical student satisfaction and test performance with a simulation-based emergency medicine curriculum: a randomized controlled trial. Ann Emerg Med 2009; 54(5): 684–91.
14. Botezatu M, Hult H, Tessma MK, Fors U: Virtual patient simulation: knowledge gain or knowledge loss? Med Teach 2010; 32(7): 562–8.
15. Satava RM: Accomplishments and challenges of surgical simulation: dawning of the next-generation surgical education. Surg Endosc 2010; 15(3): 232–41.
16. Gaba DM, DeAnda A: A comprehensive anesthesia simulation environment: re-creating the operating room for research and training. Anesthesiology 1988; 69(3): 387–94.
17. Alinier G, Hunt WB, Gordon R: Determining the value of simulation in nurse education: study design and initial results. Nurse Educ Pract 2004; 4(3): 200–7.
18. Radhakrishnan K, Roche JP, Cunningham H: Measuring clinical practice parameters with human patient simulation: a pilot study. Int J Nurs Educ Scholarsh 2007; 4: Article8.
19. Cendan JC, Johnson TR: Enhancing learning through optimal sequencing of web-based and manikin simulators to teach shock physiology in the medical curriculum. Adv Physiol Educ 2011; 35(4): 402–7.
20. Crochet P, Aggarwal R, Dubb SS, et al: Deliberate practice on a virtual reality laparoscopic simulator enhances the quality of surgical technical skills. Ann Surg 2011; 253(6): 1216–22.
21. Lee TL, Son JH, Chandra V, Lilo E, Dalman RL: Long-term impact of a preclinical endovascular skills course on medical student career choices. J Vasc Surg 2011; 54(4): 1193–200.
22. Privett B, Greenlee E, Rogers G, Oetting TA: Construct validity of a surgical simulator as a valid model for capsulorhexis training. J Cataract Refract Surg 2010; 36(11): 1835–8.
23. Coles TR, John NW: The effectiveness of commercial haptic devices for use in virtual needle insertion training simulations. In: 2010 Third International Conference on Advances in Computer-Human Interactions

(ACHI 2010), pp 148–53. Piscataway, NJ, The Institute of Electronic and Electrical Engineers, 2010. Available at http://www.computer.org/csdl/proceedings/achi/2010/3957/00/3957a148-abs.html; accessed May 20, 2013.
24. Barsuk JH, McGaghie WC, Cohen ER, O'Leary KJ, Wayne DB: Simulation-based mastery learning reduces complications during central venous catheter insertion in a medical intensive care unit. Crit Care Med 2009; 37(10): 2697–701.
25. Holzinger A, Kickmeier-Rusta MD, Wassertheurera S, Hessinger M: Learning performance with interactive simulations in medical education: lessons learned from results of learning complex physiological models with the HAEMOdynamics SIMulator. Comput Educ 2009; 52(2): 292–301.
26. Simon HA: Administrative Behavior. Ed 4. New York, Free Press/Simon & Schuster, 1997.
27. Becker GS: The Economic Approach to Human Behavior. Chicago, IL, University of Chicago Press, 1976.
28. Keeney RL, Raiffa H: Decisions With Multiple Objectives: Preferences and Value Tradeoffs. Cambridge, UK, Cambridge University Press, 1976.
29. Von Neumann J, Morgenstern O: Theory of Games and Economic Behavior. Princeton, NJ, Princeton University Press, 1944.
30. Mishan EJ, Quah E: Cost-Benefit Analysis. London, UK, Routledge, 2007.
31. Levin HM, McEwan PJ: Cost-Effectiveness Analysis. Thousand Oaks, CA, Sage Publications, 2001.
32. Fletcher JD: Education and training technology in the military. Science 2009; 323(5910): 72–5.
33. Kearsley G: Costs, Benefits, and Productivity in Training Systems. Reading, MA, Addison-Wesley, 1982.
34. Fitzpatrick JL, Sanders JR, Worthen BR: Program Evaluation: Alternative Approaches and Practical Guidelines. Ed 3. New York, Allyn & Bacon, 2003.
35. Phillips JJ: Return on Investment in Training and Performance Improvement Programs. Ed 2. Oxford, UK, Butterworth-Heinemann, 2003.
36. McDavid JC, Ingleson LRL: Program Evaluation and Performance Measurement. Thousand Oaks, CA, Sage Publications, 2006.
37. Thompson MS: Benefit-Cost Analysis for Program Evaluation. Beverly Hills, CA, Sage Publications, 1980.
38. Fletcher JD, Chatham RE: Measuring return on investment in military training and human performance. In: Human Performance Enhancements in High-Risk Environments, pp 106–28. Edited by O'Connor PE, Cohn JV. Santa Barbara, CA, Praeger/ABC-CLIO, 2010.
39. Fletcher JD, Hawley DE, Piele PK: Costs, effects, and utility of microcomputer assisted instruction in the classroom. Am Educ Res J 1990; 27(4): 783–806.
40. Jamison DT, Fletcher JD, Suppes P, Atkinson RC: Cost and performance of computer-assisted instruction for education of disadvantaged children. In: Education as an Industry, pp 201–40. Edited by Froomkin JT, Jamison DT, Radner R. Cambridge, MA, Ballinger Publishing, 1976.
41. Levin HM, Glass GV, Meister GR: Cost-effectiveness of computer-assisted instruction. Eval Rev 1987; 11(1): 50–72.
42. American College of Physicians (ACP): Primer on cost-effectiveness analysis. Effective Clin Pract 2000; 3(5): 253–5. Available at http://www.vaoutcomes.org/downloads/Cost_Effectiveness_Analysis.pdf; accessed April 10, 2012.
43. Tsai SL, Chen MB, Yin TJ: A comparison of the cost-effectiveness of hospital-based home care with that of a conventional outpatient follow-up for patients with mental illness. J Nurs Res 2005; 13(3): 165–73.
44. Drummond M, McGuire A: Economic Evaluation in Health Care: Merging Theory With Practice. Oxford, UK, Oxford University Press, 2001.
45. Vanhook PM: Cost-utility analysis: a method of quantifying the value of registered nurses. OJIN 2007; 12(3).
46. Petitti DB: Meta-Analysis, Decision Analysis, and Cost-Effectiveness Analysis: Methods for Quantitative Synthesis in Medicine. New York, Oxford University Press, 2000.
47. Roscoe SN, Williges BH: Measurement of transfer of training. In: Aviation Psychology, pp 182–93. Edited by SN Roscoe. Ames, IA, Iowa State University Press, 1980.
48. Taylor HL, Lintern G, Hulin CL, Talleur DA, Emanuel T, Phillips SI: Transfer of training effectiveness of a personal computer aviation training device. Int J Aviat Psychol 1999; 9(4): 319–35.
49. Orlansky J, Knapp MI, String J: Operating Costs of Aircraft and Flight Simulators (IDA Paper P-1733). Alexandria, VA, Institute for Defense Analyses, 1984. (DTIC AD-A144241). Available at http://www.dtic.mil/dtic/tr/fulltext/u2/a144241.pdf; accessed May 20, 2013.
50. Povenmire H, Roscoe SN: Incremental transfer effectiveness of a ground-based general aviation trainer. Hum Factors 1973; 15(6): 534–42.
51. Morrison JE, Holding DH: Designing a Gunnery Training Strategy (Technical Report 899). Alexandria, VA, US Army Research Institute for the Behavioral and Social Sciences, 1990. (DTIC AD-A226129). Available at http://www.dtic.mil/cgi-bin/GetTRDoc?AD=ADA226129; accessed May 20, 2013.
52. Taylor HL, Talleur DA, Emanuel TW Jr., Rantanen EM, Bradshaw GL, Phillips SI: Incremental Training Effectiveness of Personal Computer Aviation Training Devices (PCATD) Used for Instrument Training (Final Technical Report ARL-02-5/NASA-02-3). Savoy, IL, Aviation Research Laboratory Institute of Aviation, University of Illinois at Urbana-Champaign, 2002. Available at http://www.aviation.illinois.edu/avimain/papers/research/pub_pdfs/techreports/02-05.pdf; accessed May 20, 2013.
53. Holman GL: Training Effectiveness of the CH-47 Flight Simulator. Research Report 1209. Alexandria, VA, US Army Research Institute for the Behavioral and Social Sciences, 1979. (DTIC AD-A072317). Available at http://www.dtic.mil/cgi-bin/GetTRDoc?AD=ADA072317; accessed May 20, 2013.
54. Thorndike EL: The influence of first year Latin upon the ability to read English. School Soc 1923; 17: 165–8.
55. Thorndike EL: The Fundamentals of Learning. Ed 1. New York, Teachers College, 1932.
56. Cree VE, Macaulay C: Transfer of Learning in Professional and Vocational Education. London, UK, Routledge, 2000.
57. Perkins DN, Salomon G: Transfer of Learning (Contribution to the International Encyclopedia of Education). Ed 2. Oxford, England, Pergamon Press, 1992.
58. Cohn J, Fletcher JD; What is a pound of training worth? Frameworks and practical examples for assessing return on investment in training. In: Proceedings of the InterService/Industry Training, Simulation and Education Conference (I/ITSEC) 2010. Arlington, VA, National Training and Simulation Association, 2010. Available at http://ntsa.metapress.com/link.asp?id=x81460715365274p; accessed May 20, 2013.
59. Jones MB, Kennedy RS: Isoperformance curves in applied psychology. Hum Factors 1996; 38(1): 167–82.
60. de Weck OL, Jones MB: Isoperformance: analysis and design of complex systems with desired outcomes. Systems Eng 2006; 9(1): 45–61.
61. Bickley WR: Training Device Effectiveness: Formulation and Evaluation of a Methodology. Research Report 1291. Alexandria, VA, US Army Research Institute for the Behavioral and Social Sciences, 1980. (DTIC AD-A122777). Available at http://www.dtic.mil/dtic/tr/fulltext/u2/a122777.pdf; accessed May 20, 2013.
62. Carter G, Trollip S: A constrained maximization extension to incremental transfer effectiveness, or, how to mix your training technologies. Hum Factors 1980; 22(2): 141–52.