Adaptive Instructional Systems and Digital Tutoring

J. D. Fletcher(✉)

Institute for Defense Analyses, Alexandria, VA, USA
fletcher@ida.org

Abstract. This paper discusses the application of machine (or artificial) intelligence intended to accelerate the development of technical expertise by novices. It provides evidence from various sources of the instructional effectiveness of this approach. It also discusses and provides examples of the monetary and operational value of individualizing instruction beyond that obtainable in classroom instruction and points to a future where learning and/or performance aiding is provided anytime, anywhere through a combination of human guidance and tutorial dialogues with computers.

Keywords: Acceleration of expertise · Adaptive instructional system · Computer-assisted instruction (CAI) · Defense Advanced Research Projects Agency (DARPA) · Digital Tutoring · Effectiveness evaluation · Individualization · Intelligent tutoring system

1 Introduction

Efforts have been underway for a number of years to accelerate the acquisition of expertise by developing computer-assisted, adaptive instructional systems that provide the benefits and capabilities presently found in human-based, one-on-one instructional dialogues. It reviews the history, characteristics, and benefits of using computers to provide individualized, tutorial instruction using these dialogues, the design and assessment of a specific digital tutor in response to a Defense Advanced Research Projects Agency (DARPA) challenge, and implications of that effort and its results for adaptive instructional systems in general.

1.1 The Need for Adaptive Instruction

In 1890, William James [1] stated as his First Principle of Perception that: "Whilst part of what we perceive comes through our senses from the object before us, another part (and it may be the larger part) always comes out of our mind" (p. 747, 1890/1950). If individual learners differ, as they invariably do, then it is likely that their perceptions, cognition, and learning differ. This observation leads to Thorndike's [2] frustration with classroom instruction and his assertion in 1906 that "The practical consequence of the fact of individual differences is that every general law of teaching has to be applied with consideration of the particular person" (p. 83).

R. A. Sottilare and J. Schwarz (Eds.): HCII 2019, LNCS 11597, pp. 615–633, 2019.
https://doi.org/10.1007/978-3-030-22341-0_48

These views continue to be supported by empirical research on the extent of individual differences that teachers and instructors must contend with in classroom-based instruction. Early on, Suppes, Fletcher and Zanotti [3, 4] found 1:4 as the ratio in time needed by fastest to slowest learners in elementary school mathematics. Later, Gettinger [5] reported similar differences in time to learn of 1:3 and 1:5. These and other results suggest that some individuals in a classroom are being denied valuable learning opportunities and the advanced competencies they could be acquiring while others, equally deserving of learning, struggle to keep up. Gettinger emphasized, in accord with Carroll [6], that this difference is not due solely to inborn native ability, but, more precisely, to what they both described as individual differences in learning ability.

Along with learning ability, another often noted and accepted source of difference in time to learn is prior learning [7]. Because differences in prior knowledge acquired by individuals increase through time and life experience, it is likely that differences in learning rate similarly increase with the age and the varied experiences of individuals as they pursue higher education and workforce training. Overall, adapting learning to individual differences in background, temperament, ability, and prior knowledge appears to be an intransigent and continuing challenge for classroom instruction at all levels.

The general problem of individual differences can be eased by classroom practices and heroic efforts of classroom teachers, but only partially. Despite its short term economic advantages, classroom-based instruction, with its difficulty in attending to the specific needs and capabilities of individual learners, presents an unavoidable impediment to instructional efficiency and effectiveness. Based on his own and his students' research, Bloom [8] famously claimed a learning increase of two standard deviations in learning from tutoring (one instructor working with one learner) compared to classroom instruction – a difference that (roughly and on average) would increase 50th percentile learners to the 98th percentile.

Although discussion about Bloom's empirical findings continues, subsequent research supports the substantial superiority of individual tutoring over classroom instruction [9–12]. Why then do we not provide an Aristotle for every Alexander and a Mark Hopkins for the rest of us? The answer, of course, is that, except for very complex and critical activities (e.g., surgery, airplane piloting), we cannot afford it.

But we can afford computers. Following the development of writing, which made learning portable, and then books, which made learning both portable and (eventually) affordable, we may be on the verge of a third revolution in the teaching-learning process – the use of computer technology to provide universal, adaptive, on-demand, tutorial instruction. Full natural language, with its use of metaphors, similes, allusions, slang, and other peculiarities, may remain beyond the capabilities of computers for some time, but their understanding of natural language has been sufficient to support a wide range of tutorial interactions for some time [13, 14].

This possibility suggests a future in which personal computer-based devices (e.g., laptops, telephones) provide adaptive instruction, performance aiding, and decision support in the form of tutorial dialogues at any time and practically anywhere. Necessary and fully up to date subject matter information (updates and all) need not be stored locally. It can be collected in real time and as required from the global information grid (the “cloud”) and adapted to the background, needs, evolving capabilities, and interests of individual learners [15].

1.2 Characteristics of Adaptive Instruction

Today, and according to the US Department of Education [16] adaptive instructional systems cluster around one of three strategic approaches: differentiation, which adapts instructional approaches for different groups of learners; personalization, which adapts the topics and objectives of instruction to learners’ specific needs, interests, and/or preferences; and individualization, which adapts instruction, but not its topics and objectives, to the abilities, prior learning, and learning progress of each learner as an individual.

Choosing which strategy to apply in devising an adaptive instructional system may be keyed to the overall purpose of the instruction, specifically, whether it is intended for education or training. For instance, education must adequately prepare learners for unknown futures, whereas training must prepare learners for a known and, to an extent, understood future, i.e., specific tasks and occupations. For this reason, individualization, with its focus on attaining specific, prescribed learning objectives, may be better applied in training, especially technical training, while personalization, with its adjustment of objectives to the learner, may be a better choice for education and education institutions. Differentiation, with its adaptations for different groups of learners rather than individuals may be closer to personalization than individualization. At present, digital tutoring seems more focused than differentiation or personalization on adaptive instruction as training, with its requirement that individual learners achieve specific learning objectives at requisite standards and conditions.

When it comes to adapting instruction, the differences between training and education are neither rigid, pure, nor absolute. Training and education may be viewed as opposite ends of a continuum of instruction. Most training includes elements of education, and most education includes elements of training. Both physicists and electronic technicians learn Ohm’s Law and algebra through education, but both may need to be trained in the skills needed to operate an oscilloscope. Both surgeons and Boatswain’s mates must acquire skill in tying knots, but both must understand when and why to use them in their work. Nonetheless, both education and training are fundamentally concerned with learning and cognition. Techniques and findings from either may be as relevant and applicable for the opposite end of the continuum as they are to their own. Increased attention of scholars and researchers to both ends of this continuum seems in order, if not overdue. Given this context, we may turn to the use of computers to adapt instruction to learners.

1.3 Computers and Adaptive Instruction

Like many innovations (e.g., horseless carriages, wireless telegraph), computer-assisted instruction (CAI) began by piling an existing technology, in this case textbook based programmed learning, onto a new technology. Programmed learning applied in CAI is based on processes and frames such as the one shown in Fig. 1. Typically, it applies

Keller's [17] Personalized System of Instruction to determine which set of programmed learning frames a learner should receive. Then, in accord with Crowder's [18] Intrinsic Programming – as opposed to Skinner's [19] Extrinsic Programming, it presents instruction consisting of expository text and frames, to provide the next steps for instruction depending on the learner's response.

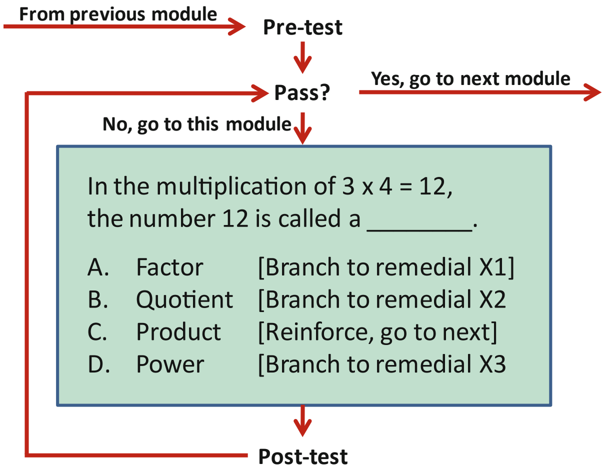

Fig. 1. A representative intrinsic programming frame

It is relatively inexpensive to program computers for these frames, and the approach remains in wide use. Reviews found it to be modestly superior to classroom learning, reporting moderate effect sizes in the area of 0.43, thereby indicating improvements in learning by 50th percentile learners (roughly and on average) to the 67th percentile [20, 21].

However, frame oriented instruction requires considerable human "authoring" effort and expense because the frames must anticipate and provide for every possible state of the learner and the instruction, which was early shown to be impossible – even for something as rudimentary as second grade arithmetic [22]. A more adaptive and effective approach was evidently needed – an approach in which the state of the learner and the instruction could be dynamically generated, modeled, and provided by computer. This possibility motivated the Department of Defense to support research and development, very early, for intelligent tutoring systems [23], identified here as digital tutors.

2 Digital Tutoring

With support in the mid-1960's from the Office of Naval Research, Wallace Feurzeig [14] determined that computers could and should do more than simply mimic programmed textbooks. He developed a computer language (MENTOR) and a program (also called MENTOR) to prepare learners to perform medical examinations. Based on initial development of the Mentor language, semantic networks, Quillian's "language comprehender" [24], and his own work with SCHOLAR, Carbonell [25] identified two definitive features that distinguish digital tutors from standard CAI:

- Information system structures. These are described as dynamically generated information structures provided in place of pre-programmed, frame-oriented approaches. Generative, information structures, such as those based on ontologies, concept maps, natural-language understanding, and one-on-one tutorial capabilities, relieve developers from the need to anticipate every state that might exist for an individual learner and the instructional system. Modeling these states and dealing with them were thereby assigned, as much as possible, to the computer.
- Mixed-initiative tutorial dialogue. Digital tutors allow either the computer or the learner to take the initiative in asking questions, posing problems, or initiating lines of inquiry. Students can ask questions at various levels of understanding or abstraction, and tutorial responses by the computer can be generated and tailored for the level of learning the student has achieved as well as the context within which the responses occur.

Dialogue		Comments
Student	**What is the base emitter voltage of the voltage limiting transistor?**	This exchange is straightforward question and answer. SOPHIE introduced a fault into the simulated power supply and the student must locate and repair it.
SOPHIE	**The base emitter voltage of Q5 is -.037 volts.**	
Student	**What about for the current limiting transistor?**	SOPHIE must be able to parse the dialogue sufficiently to know that the student is asking for the base emitter voltage of transistor Q8. The student's question derives from the context established by SOPHIE.
SOPHIE	**The base emitter voltage of Q8 is .349 volts.**	
Student	**What should it be?**	Again, the interaction depends on SOPHIE's understanding of context. It also requires SOPHIE to simulate both the faulted power supply and one that is operating correctly.
SOPHIE	**The base emitter voltage of Q8 is .349 volts.**	
Student	**Replace Q8.**	This interaction illustrates a clear step beyond what Brown et al. considered knowledgeable to one they considered intelligent. SOPHIE has knowledgeably parsed both dialogue and the student's emerging solution path, modeled the student's troubleshooting hypotheses, determined that they are incorrect, is capturing the dialogue initiative from the student, and is undertaking a series of tutorial interactions intended to guide the student back to a correct solution path.
SOPHIE	**I am going to ask you some questions about how Q8 is faulted. Are any junctions shorted?**	

Fig. 2. Sample tutorial dialogue with SOPHIE [13]

These distinctions were based on what Carbonell [8] called Information Structure Oriented Instruction as opposed to the Ad-Hoc Frame Oriented instruction that used programmed learning techniques. This approach was later the basis for developing, among others, the SOPHIE system to train electronic technicians as illustrated in Fig. 2. The comments in Fig. 2 annotate a dialogue recorded between a learner troubleshooting an electronic power supply and a digital tutor (SOPHIE). As well as providing advanced natural language techniques that demonstrate SOPHIE's adaptive dialogue capabilities, the final exchange shows SOPHIE's mixed initiative capabilities. Like a good tutor, it knew when to take the initiative from the learner and how to ask questions that guide the learner to a correct solution path.

As suggested from the 1970s on [e.g., 26], digital tutors typically apply three explicit models used to adapt instruction: (1) a model of the subject matter including the knowledge and skills to be acquired; (2) a dynamically evolving model of each learner's understanding and acquisition of the desired knowledge and skills; and (3) a model of instructional techniques, i.e., tutorial strategies, that may be used by a specific learner to develop knowledge and skills derived from the first two models.

Later, an additional model (4) was added to cover communication between the learner and the instructional system [27, 28]. The first, third, and fourth of these models may be devised at the beginning of the instruction. The third and fourth together provide a theme on which tutorial interactions might be improvised in real time based on emerging characteristics and preferences of the learner. The second model must evolve dynamically with the student. It must be generated, adapted, and revised in real time, preferably in a "stealthy" manner [29], with a minimum of explicit testing. It provides a foundation for adapting instruction to the learner.

2.1 Effectiveness of Digital Tutoring

Assessments of digital tutors generally report greater statistical effectiveness compared to other instructional approaches. However, among them, and in other research as well, there are findings that are statistically significant, but of minor instructional value and/or return on investment time and cost. For that reason, researchers increasingly use effect sizes to indicate the practical value of their findings.

Effect sizes report differences between experimental groups in terms of standard deviations. Discussion about the proper calculation of effect sizes and their interpretation continues [30, 31]. These differences are notable, and worthy of consideration, but their calculation is not as much an unsettled issue as their interpretation, which may be peculiar to the individuals or organizations using them for decision-making. Table 1 suggests some interpretations of effect sizes. It is based on comments by Cohen [32], the DoEd What Works Clearinghouse [16], and Bloom [8] and suggests a guide for interpreting effect sizes, or at least those reported here.

Earlier (1994), evaluations of frame-based CAI found that it produces more learning than classroom instruction with standard deviations in the area of 0.33 [33, 34]. This result would be viewed as a small, but appreciable improvement in instruction, given the interpretations suggested in Table 1. However, it is appreciably less than reviews of digital tutoring, which appears to be superior in providing learning than either classroom instruction or frame-based CAI.

Table 1. Overview of effect size

Effect Size	Suggested designation[a]
ES < 0.25	Negligible[b]
0.25 < ES < 0.40	Small
0.40 < ES < 0.60	Moderate
0.60 < ES < 0.80	Large
ES > 1.00	Very large
ES > 2.00	Bloom's challenge[c]

Notes: [a]Extended from suggestions by Cohen [32]; [b]What Works Clearinghouse [16]; [c]Bloom [8]

Meta-analyses of digital tutoring have been performed by VanLehn [35] and Kulik and Fletcher [36]. VanLehn reviewed 27 studies of digital tutoring and found that they averaged a moderate effect size of 0.59. However, in further investigation of his data, he found an average effect size of 0.40 for sub-step-based tutoring compared to a large effect size of 0.76 for step-based tutoring. In other words, learning by 50th percentile learners would improve (roughly and on average) to the 66th percentile under fine-grained tutoring – about the same as frame-based learning – but improve (roughly and on average) to the 78th percentile under more general, less specific tutorial interactions. Additional research may better account for this finding, which may have been due to the need for students to reflect more carefully and thoroughly on results of step-based tutoring than those of sub-step-based tutoring.

An extensive and more recent analysis by Kulik and Fletcher [36] found a large effect size of 0.66 for 50 digital tutors, with data ranging from –0.34 to 3.18 (after Winsorizing for outliers) in effect size – an overall result between VanLehn's analysis of sub-step and step-based tutoring, but appreciably closer to the latter than the former. In any case, these findings, which used a more precise definition of digital tutoring than some earlier analyses, suggest substantial learning improvements over both classroom instruction and applications of frame-based programmed learning techniques in CAI.

2.2 Design and Development of the DARPA Digital Tutor

Given this context, the design, development, and two recent assessments of a digital tutor developed by the Defense Advanced Research Projects Agency (DARPA) for the US Navy deserve attention. Development of this technology has been proceeding steadily since Feurzieg's MENTOR [14]. However, DARPA's mission is to leap ahead and develop high payoff research that is too risky and expensive to be considered by the military Service laboratories. An example is the development of cigarette package sized devices to replace suitcase-sized systems for determining locations on earth using Global Positioning System satellites.

In education and training, DARPA's intent was to substantially accelerate acquisition of expertise, well beyond journeyman levels, by learners starting with little, if any, prior knowledge or training in the subject area. The subject matter chosen was

Information Systems Technology, which is abbreviated by the Navy as IT in reference both to the technology and to individuals with this occupational specialty.

Design and development of the Tutor was basically a matter of identifying high-quality tutorial ingredients and applying them in proportions determined by system-atic empirical testing. It was pragmatic and not focused on verifying any particular theory of learning, cognition, and/or instruction. The developers had acquired years of experience studying one-on-one tutoring by humans. Findings from that work provided initial approaches for designing the DARPA Tutor.

DARPA funding allowed extensive assessments in devising the Digital Tutor [37–39]. Early assessments of tutoring techniques used IT novices of about the same age, education level, and capabilities as novice Sailors assigned to the Navy IT school. The Tutor was developed step by step in sessions with detailed video and additional voice recordings to determine what was effective and what was not. These sessions were repeated as needed to develop instructional activities and interactions that reliably accomplished targeted learning objectives. This iterative approach provided the foundation for a problem-based learning environment allowing computer development of subject matter and individualized information structures that were functionally similar to those of human tutors with expertise in both the subject matter and in one-on-one instruction.

Experts in specific IT topics were contacted based on their knowledge, publications, and reputation. These experts were auditioned in 30-min IT tutorial sessions with learners who were representative of new Navy sailors. The intention was to base (or "clone") the Digital Tutor using the practices of individuals who were expert in both an IT topic and in one-on-one tutoring.

These sessions helped identify and select 24 tutors who were experts in requisite IT topics for use in designing the Digital Tutor. The tutors then tutored 15 IT qualified sailors who were newly graduated from recruit training and chosen at random. The sailors were tutored one-on-one by these experts for 16 weeks to prepare them for IT careers in the Navy. Every session in this tutoring was again captured in video. These sessions, which were extensively reviewed and assessed, served as the basis for the tutorial instruction provided by the Digital Tutor. This analytic work, including further trial and analyses, continued as the Tutor was developed. The Tutor employs the following prescriptive procedures:

- Promote reflection by eliciting learner explanations of what went well and what did not;
- Probe vague and incomplete responses;
- Allow learners to discover careless errors but assist learners in correcting errors arising from lack of knowledge or misconceptions;
- Never articulate a misconception, provide the correct answer, or give a direct hint;
- In the case of a learner impasse, review knowledge and skills already successfully demonstrated by the learner and probe for why they might or might not be relevant to the current problem; and
- Require logical, causal, and/or goal-oriented reasoning in reviewing or querying both incorrect and correct actions taken by the learner to solve problems.

The Tutor used information structures to:

- Model the subject matter;
- Generate evolving models of the learner;
- Generate, adapt, and assign problems to maximize progress of individual learners;
- Engage in tutorial exchanges that shadow, assess, and guide learners' problem solving;
- Ensure that learners reflect on and understand deeper issues and concepts illustrated by the problems.

Operationally, the design of the Digital Tutor emphasizes:

- Active, constant interaction with learners – which fostered the "flow" that is found in computer-based games [40].
- Capturing in digital form the processes and practices of one-on-one tutoring;
- Requiring problem solving in authentic environments – leaners used actual Navy systems, not simulations of these systems. Problem solving was not based on copying the problem solving paths of experts. The tutor was expected to help learners follow whatever path they chose to troubleshoot and solve problems.
- Continual, diagnostic assessment of individual learner progress;
- Focus on higher order concepts underlying problem solving processes and solutions;
- Integration of human mentors.

The presence of experienced Navy ITs as mentors was essential for this training. They resolved difficulties in human-computer communication, managed the study halls, and, especially, provided examples of Navy bearing and culture for the novice sailors. "Sea stories" might be viewed as little more than entertainment, but, as with Army "War Stories" and Air Force "Air Stories", few activities are as effective and important as these stories in providing civilians with the esprit de corps and culture needed to prepare them for military service.

Mirroring its development strategy, the Tutor's instructional approach is spiral. It presents conceptual material selected for individual learners by the tutor. This material is immediately applied in solving problems intended to be comprehensive and authentic. Learners interact directly with US Navy IT systems, not simulations, while the Tutor observes, tracks, and models their progress and solution paths. Tutoring tactics developed for the Tutor were the following:

- Promote learner reflection and abstraction by:
  - Prompting for antecedents, explanations, consequences, or implications of answers.
  - Questioning answers, both right and wrong.
  - Probing vague or incomplete explanations and other responses by the learner.
- Review knowledge and skills when the learner reaches an impasse or displays a misconception by asking why something did or did not happen.
- Avoid providing a correct answer, providing a direct hint, or articulating a misconception.

- Sequence instruction to pose problems that are tailored and selected to optimize each learner's progress.
- Require logical, causal, or goal-oriented reasoning in reviewing or querying steps taken by the learner to solve problems.
- Refocus the dialogue if the learner's responses suggest absent or misunderstood concepts that should have been mastered.
- If a learner makes a careless error in applying a concept already mastered, allow problem-solving to continue until the learner discovers it.
- Verify learner understanding of any didactic material before proceeding.

A daily schedule for instruction consisted of 6 h using the Tutor followed by a two-hour study hall, which was proctored by one of the Navy instructors assigned to the school. It involved discussion and reflection on material presented during the day. At the end of the week, one of the senior designers of the Tutor would attend to participate in the discussion, address particularly difficult issues that the learners encountered during week, and, in return, gain insight into what the Tutor was doing well and not well.

2.3 Effectiveness of the DARPA Digital Tutor for the Active Navy

Navy Assessment
After the training was about half finished, IT knowledge of the human tutored sailors was assessed by a paper-and-pencil test prepared by Navy instructors. It included multiple choice, network diagram, and essay questions answered by the 15 human-tutored and by 17 classroom instructed sailors. The tutored sailors averaged 77.7 points compared to 39.7 points for the classroom training sailors on this test, which indicated an effect size of 2.48 exceeding Bloom's 2.00 target in their favor [39].

Four other assessments of the Digital Tutor were performed with different sailors participating at progressive stages of the Tutor's development: 4 weeks, 7 weeks, 10 weeks, and finally 16 weeks [37, 39].

The first assessment of the Tutor compared the IT knowledge of 20 new sailors, who had completed the first 4 weeks of Digital Tutor training then available, with that of 31 sailors who had graduated from approximately 10 weeks of standard classroom training and with that of 10 Navy IT instructors. This study the found an effect size of 2.81 in favor of the 4-week Tutor students over the students who had graduated from the 10 week IT course and an effect size of 1.32 in their favor compared to their Navy instructors. These differences were also statistically significant ($p < 0.05$) [39].

The next assessment compared both the IT trouble shooting ability and IT knowledge of 20 new sailors, who had completed the 7 weeks of the Digital Tutor training then available, with that of 20 sailors who had graduated from a newly revised 19-week IT classroom and laboratory training course and with that of 10 instructors who only took the knowledge test. The IT trouble shooting effect size favoring the 7-week Tutor sailors' troubleshooting skill over that of the 19-week classroom and laboratory sailors was 1.86. The IT knowledge difference effect size favoring the 7-week Tutor sailors over the 19-week classroom and laboratory sailors was 1.91 and it was 1.31 in comparison with their instructors. Again these differences were statistically significant ($p < 0.05$) [37, 39].

A final assessment was performed after another representative group of 12 sailors had completed training with the final 16-week version of the Tutor. The DARPA challenge was to produce in 16 weeks (the usual time for ab initio IT training) sailors who would be superior in skill and knowledge to (a) other novice sailors trained using conventional classroom and laboratory practice, and (b) ITs with many years of experience in the Fleet [39].

The assessment involved new sailors who completed IT training with the DARPA Digital Tutor, other new sailors trained for 35 weeks using the Navy's classroom based Information Technology Training Continuum (ITTC), and senior ITs with an average of 9.2 years of Fleet experience. As with all the assessments, sailors who had just finished recruit training were assigned at random to the two training groups (DT and ITTC standard classroom training with laboratory experience). The Fleet ITs were chosen as the "go to" ITs from ships on shore duty in San Diego and Oak Harbor, Washington.

There were 12 ITs in each group. Repeated measures were used because of the small sample sizes – 14 h of IT trouble shooting skill testing, 4 h of written (mostly short answer) knowledge testing were used in the assessments. Other tests such as oral examination by experienced ITs, development and design of IT systems according to typical specifications, and ability to ensure security of an IT system were also applied.

IT troubleshooting in response to trouble tickets was the most important component of the training in preparing these novice sailors for their Navy IT occupation. It was intended to resemble Fleet IT requirements as closely as possible. Navy "Trouble Tickets" which had been submitted from the Fleet for shore-based assistance were presented as problems to be solved by 3-member teams with a specified time for solution. Results of the Troubleshooting testing are shown in Fig. 3. Notably neither the ITTC nor the Fleet teams attempted to solve the "Very Hard" problems.

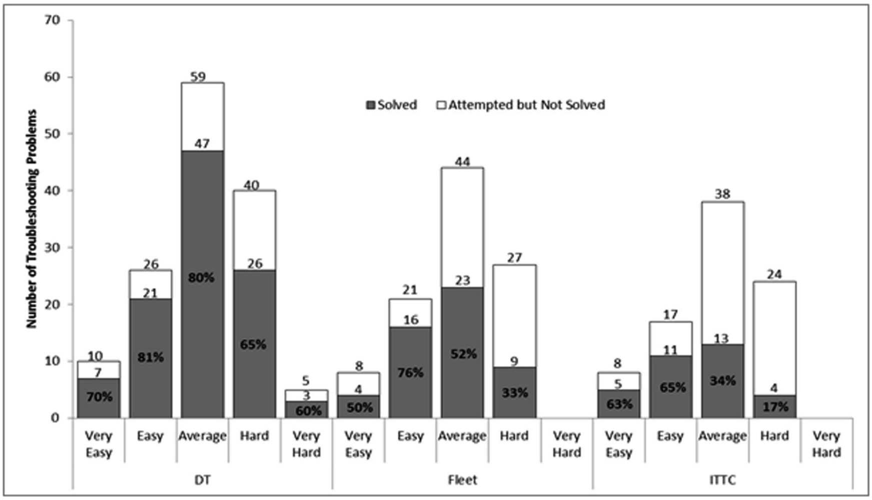

Fig. 3. Troubleshooting problems solved by DT, Fleet, and ITTC teams

Troubleshooting capability was the primary focus of the assessment because it best indicated how well the new ITs were prepared to do their jobs in the Fleet. Acquisition of IT knowledge was also of interest and found to account for about 40% of individuals' troubleshooting scores. This is an appreciable amount and it is of interest, but performance in IT troubleshooting is the main concern of the Navy. More description of these assessments along with additional data, testing, and findings is provided by [39].

Veterans Assessment

An approximate replication of the above assessment was provided by assessing an 18- week version of the Tutor used to train 100 military veterans [38]. The course was extended by 2 weeks to assist the veterans in adapting and applying to civilian technical workplaces. As Table 2 shows, most of the veterans were unemployed before taking this course. There were no academic dropouts from the course, which was completed by 97 of the veterans. Fourteen of the veterans chose to seek higher education rather than apply directly for employment. Another 6 veterans did not reply to requests for post-training information. All 77 of the graduates who sought employment were hired. Their average annual salary was $73,000, which, at the time, was equivalent to civilian employment intended for IT technicians with 3–5 years of IT experience [41]. Most received early bonuses and promotions.

Table 2. Characteristics of 101 Veterans[a] accepted for digital tutor IT training

Average years of separation from military service	5.20
Avg age	30.5
Married	30
Armed forces qualification test	87.1
Full time employment	11
Part time employment	45
Prior civilain IT instruction	8
High school/GED degree	45
AA degree	11
BA/BS	44
Other	1
Prior Military IT Instruction	4

Note: [a]One veteran dropped out before beginning the course and was replaced.

Results from return on investment analysis are shown in Fig. 4. It shows that monetary return to the government over a 20-year period is appreciable for all monetary support provided to veterans. However, the return is much greater for the 18-week digital tutoring program than government support for a either a 2 or 4 year degree and even for veterans who completed a program of education after receiving no monetary support from the Veterans Administration [38].

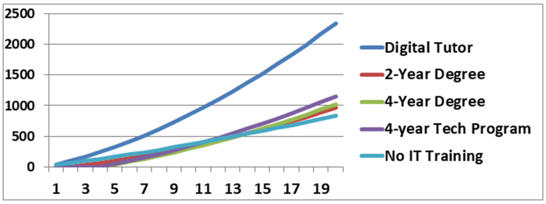

Fig. 4. Monetary Return to US Government Per Individual from Support Provided for Education and/or Training ($000)

2.4 Summary Comments on Digital Tutoring to Provide Adaptive Instruction

Assessments of the DARPA Digital Tutor suggest a number of possibilities and issues, four of which are discussed here. Others will doubtless occur to readers.

Acceleration of Expertise
The value of technical expertise is as evident from empirical research as it is from random observation [42, 43]. However, the years of experience and practice needed to develop technical expertise increase its cost and limit its supply. Empirical demonstrations that the time to develop technical expertise can be compressed from years into months are few, but extant.

For instance, the Sherlock project [44–46] prepared technicians to solve complex problems occurring in a test stand used to troubleshoot components of Air Force avionics systems. Assessments found that 20–25 h of Sherlock training produced about the same improvement in performing difficult and rarely occurring diagnostic tasks as 4 years of on-job experience [44]. Their approach presaged that of DARPA's Tutor in assuring that intensive technical learning was always followed by guided reflection on what worked and what did not.

Other evidence was provided by IMAT, the Navy's Interactive Multi-Sensor Analysis Trainer [47]. This system focused on what the authors described as 'incredibly complex tasks'. They describe these as broad, multifaceted, abstract, co-dependent, and nonlinear tasks that require a large repertoire of patterns and pattern-recognition capabilities for their solution. An at-sea trial found that 2 days of training with a laptop version of IMAT increased submarine effective search area by a factor of 10.5 [48]. In effect, a submarine with IMAT-trained sonar operators could provide the sonar surveillance of 10 submarines with operators who lacked IMAT training. The operational and monetary value of this capability is substantial.

These examples, in addition to the DARPA Digital Tutor discussed here, suggest training advances of considerable value, including those that may reliably and significantly accelerate the acquisition of expertise, waiting and within our reach, but not yet in our grasp.

Return on Investment
Training and education are often viewed as expenses, not investments, which does not serve well either of them or the many who benefit from them. If designed honestly and

well[1] digital tutors are expensive to design and build. However, the monetary and operational costs of not doing so may be far greater. No quantity of digital tutors would cost more than the loss of a single submarine, or any ship that was lost because its internal network system failed. That aside, analysis [49] found that continuing to provide standard classroom IT training with its requirement for years of follow-on development and on-job training is far more expensive than the cost to design, develop, and reliably update a digital tutor. For instance and assuming that the Navy must train 2,000 ITs a year for the Fleet, the costs saved by using the 16-week Digital Tutor program to replace 16 weeks of classroom training followed by 7 years of on-job training averaged savings in discounted dollars were estimated to average $109M annually over a 12 year period [49].

As further shown by the development and delivery of the DARPA Digital Tutor to veterans, its design, development, and delivery costs were substantially less than the costs to provide standard Veterans Administration education benefits for individuals to acquire a 2-year or 4-year college degree [38]. For that matter, individuals who completed a 4-year college degree in information network technology with no support from the government returned considerably less to the government in income than ITs who completed the Digital Tutoring program was extended to 18 weeks to prepare veterans for the civilian market place. The internal rate of return on government investment that provided 18 weeks of housing, meals, and Digital Tutoring to prepare a veteran for a career in IT was estimated to be about 35 percent over a 20-year period.

In general, training and education and their consumers, might benefit significantly if, in addition to assessing their effectiveness for learning, our evaluations and assessments of instructional capabilities additionally included assessments of their likely return on investment.

Applying Digital Tutors

As discussed above, a case can be made that Digital Tutoring is more effective and suitable for instructional objectives involving conceptual learning rather than those concerning the initial rudiments of any subject. Objectives for these rudiments, such as nomenclature, common procedures, and basic procedures, are found at the low end of Bloom's often referenced hierarchy of instructional objectives [6]. Nonetheless, they are essential for learning and instruction in most, if not all, subject matter. As early CAI programs demonstrated, these rudiments are readily learned through the techniques of drill and practice.

Most successful drill and practice programs focus on discrete items such as arithmetic facts, vocabulary words, orthography, technical terms, and the like. Drill and practice is an effective, and when well done, incentivizing approach. Its promise was early demonstrated for introducing basics in subjects such as beginning reading [50] and elementary mathematics [3, 4, 51, 52]. Comments about "drill and kill" may apply to some classroom learning, but drill and practice programs presented by computer have been found to be successful and enjoyed by early learners [34, 53, 54]. Considerable data from the 1960s, 1970s, and onward have shown these rudiments can be acquired efficiently and effectively through computer-assisted drill and practice. Some

[1] "Intelligent Tutoring" has long been used as a marketing term by training developers and contractors.

drill and practice programs have applied sophisticated approaches such as statistical optimization routines to select and present items that maximize individual learning given constraints such as time available and a learner's progress [51, 55–58]. Comparisons of these early drill and practice programs with conventional classroom instruction generally found effect sizes of about 0.40 [33].

Moreover, and as shown in Table 3, digital tutoring does not do as well in preparing learners with these rudiments as does drill and practice. Tutoring appears necessary and more suited for the next step – applying subject matter rudiments to develop the abstractions and concepts needed for a deeper and more nuanced understanding of the subject matter. As effect sizes in the table suggest, this conceptual area appears to be where digital tutoring is most needed and most successful in providing a full understanding of the subject matter and applying it successfully.

Table 3. Effect sizes for four digital tutoring systems assessed for conceptual and rudimentary learning

Source	Concepts	Rudiments
Graesser et al. [59]	0.34	0.00
Koedinger et al. [60]	0.99	0.36
Person et al. [61]	0.30	0.03
VanLehn [35]	0.95	−0.08
Average (Standard Deviation)	0.65 (0.326)	0.08 (0.168)

The idea of pairing drill and practice programs with digital tutoring is becoming less heretical in digital tutoring circles. For instance, this approach is suggested by Nye et al. [62] along with the sensible caution that drill and practice may be overdone by focusing entirely on solving specific problems. Overall, rudiments may be best left to drill and practice techniques with digital tutoring brought in once the rudiments are sufficiently acquired. Reflection, which is enabled by dialogue exchanges in digital tutoring, reveals the abstract and generalizable concepts underlying problems presented and increases both retention and transfer of what is learned [45, 46, 60, 63, 64].

Team Training

An issue raised by Fletcher and Sottilare [65] concerns Digital Tutoring capabilities applied to training for teams. As Jones [66] suggested, teams differ in their degree of "teamness". That is to say that in some teams the members perform their task almost independently, passing off their contributions without much adjustment in their actions based on what other team members do. In these cases, application of Digital Tutoring seems relatively simple – they can be trained and taught in much the same manner that individuals are trained and taught. Many positions in baseball, as Jones points out, are like this. However, in some teams the interactions of at least some members depend closely on what other team members may do. Doubles tennis is a good example of this interaction. In baseball, as Jones suggests, the pitcher and the catcher may comprise a genuine team, with each responding to actions taken by the other. Whether we agree this example or not, it seems evident that teams and their members differ in the teamness of their responsibilities.

Sottilare with others [65, 67] suggests that there is a role for intelligent tutoring in team training and that it may be organized and assisted by the Generalized Intelligent Framework for Tutoring (GIFT). Work in this area is recent and continuing. It has yet to be applied and assessed in a context for team training, but, considering the intense requirements for teams and team activity in Defense, it seems likely to proceed.

3 Final Comment

Finally, no "magic sauce" or specific academic theory was used to produce the DARPA Digital Tutor. It was designed and developed by using empirical means to identify high-quality, but well-known, tutorial ingredients and applying them in proportions determined by systematic, empirical testing.

Certainly theory for instruction is essential [51, 58] but the Digital Tutor was initiated by a DARPA challenge to solve a practical problem. The approach used to develop a solution was based on performance requirements rather than an attempt to prove a theory. Like education and training, practice and theory appear to exist on a continuum, but the Tutor was more focused on solving a practical problem, than proving a theory. Its development was fundamentally eclectic and pragmatic, based on an iterative, formative evaluation approach.

The DARPA Digital Tutor may have realized a breakthrough in the technology of adaptive learning. It was a catalyst for a 2017 National Academy of Sciences, Engineering, and Medicine symposium to press for wider and more routine use of this technology in order to prepare the national technical workforce for both present and emerging challenges to the national economy and productivity. The consensus was that digital tutoring technology is essential and ready to assume this responsibility. Nonetheless, how best to move it from the laboratory into the field remains undetermined.

References

1. James, W.: Principles of Psychology, vol I. Dover Press, New York (1890/1950)
2. Thorndike, E.L.: Principles of Teaching. A. G. Seiler & Company, New York (1906)
3. Suppes, P.C., Fletcher, J.D., Zanotti, M.: Performance models of American Indian students on computer-assisted instruction in elementary mathematics. Instr. Sci. **4**, 303–313 (1975)
4. Suppes, P.C., Fletcher, J.D., Zanotti, M.: Models of individual trajectories in computer-assisted instruction for deaf students. J. Educ. Psychol. **68**, 117–127 (1976)
5. Gettinger, M.: Individual differences in time needed for learning: a review of literature. Educ. Psychol. **19**(1), 15–29 (1984). https://doi.org/10.1080/00461528409529278
6. Carroll, J.B.: Problems of measurement related to the concept of learning for mastery. Educ. Horiz. **48**, 71–80 (1970)
7. Tobias, S.: Extending Snow's conceptions of aptitudes. Contemp. Psychol. **48**, 277–279 (2003)
8. Bloom, B.: The 2 sigma problem: the search for methods of group instruction as effective as one-to-one tutoring. Educ. Res. **13**(6), 4–16 (1984)
9. Cohen, P.A., Kulik, J.A., Kulik, C.-L.C.: Educational outcomes of tutoring: a meta-analysis of findings. Am. Educ. Res. J. **19**, 237–248 (1982)

10. Evans, M., Michael, J.: One-On-One Tutoring by Humans and Computers. Lawrence Erlbaum, Mahwah (2006)
11. Graesser, A.C., D'Mello, S.K., Cade, W.: Instruction based on tutoring. In: Mayer, R.E., Alexander, P.A. (eds.) Handbook of Research on Learning and Instruction, pp. 408–426. Routledge Press, New York (2011)
12. Graesser, A.C., Moreno, K., Marineau, J., Adcock, A., Olney, A., Person, N.: The tutoring research group. AutoTutor improves deep learning of computer literacy: is it the dialog or the talking head? In: Hoppe, U., Verdejo, F., Kay, J. (eds.) Proceedings of Artificial Intelligence in Education, pp. 47–54. IOS Press, Amsterdam (2003)
13. Brown, J.S., Burton, R.R., DeKleer, J.: Pedagogical, natural language and knowledge engineering in SOPHIE I, II, and III. In: Sleeman, D., Brown, J.S. (eds.) Intelligent Tutoring Systems, pp. 227–282. Academic Press, New York (1982)
14. Feurzeig, W.: Computer systems for teaching complex concepts (BBN Report 1742). Bolt Beranek & Newman, Inc., Cambridge, MA (DTIC AD 684 831) (1969)
15. Fletcher, J.D., Tobias, S., Wisher, R.L.: Learning anytime, anywhere: advanced distributed learning and the changing face of education. Educ. Res. **36**(2), 96–102 (2007)
16. USDoEd: Transforming American education learning powered by technology. National education technology plan 2010. United States Department of Education, Office of Educational Technology, Washington, DC (2010)
17. Keller, F.S.: Goodbye, teacher. J. Appl. Behav. Anal. **1**, 79–89 (1968)
18. Crowder, N.A.: Automatic teaching by means of intrinsic programming. In: Galanter, E. (ed.) Automatic Teaching: The State of the Art, pp. 109–116. Wiley, New York (1959)
19. Skinner, B.F.: The science of learning and the art of teaching. Harv. Educ. Rev. **24**, 86–97 (1954)
20. Kulik, J.A., Cohen, P.A., Ebeling, B.J.: Effectiveness of programmed instruction in higher education: a meta-analysis of findings. Educ. Eval. Policy Anal. **2**, 51–64 (1980)
21. Kulik, C.-L.C., Schwalb, B.J., Kulik, J.A.: Programmed instruction in secondary education: a meta-analysis of evaluation findings. J. Educ. Res. **75**, 133–138 (1982)
22. Barr, A., Feigenbaum, E.: Buggy. In: Handbook of Artificial Intelligence, vol. 2, pp. 279–284. William Kaufmann, Los Altos (1982)
23. Fletcher, J.D., Rockway, M.R.: Computer-based training in the military. In: Ellis, J.A. (ed.) Military Contributions to Instructional Technology, pp. 171–222. Praeger Publishers, New York (1986)
24. Quillian, M.R.: The teachable language comprehender: a simulation program and theory of language. Commun. ACM **12**(8), 459–475 (1969)
25. Carbonell, J.R.: AI in CAI: an artificial intelligence approach to computer-assisted instruction. IEEE Trans. Man Mach. Syst. **11**, 190–202 (1970)
26. Fletcher, J.D.: Models of the learner in computer-assisted instruction. J. Comput. Based Instr. **3**, 118–126 (1975)
27. Anderson, J.R., Boyle, C.F., Corbett, A., Lewis, M.W.: Cognitive modeling and intelligent tutoring. Artif. Intell. **42**, 7–49 (1990)
28. Koedinger, K.R., Anderson, J.R., Hadley, W.H., Mark, M.A.: Intelligent tutoring goes to school in the big city. Int. J. Artif. Intell. Educ. **8**, 30–43 (1997)
29. Shute, V., Ventura, M.: Stealth Assessment: Measuring and Supporting Learning in Video Games. MIT Press, Cambridge (2013)
30. Ellis, P.D.: The Essential Guide to Effect Sizes. Cambridge University Press, Cambridge (2010)
31. Grissom, R.J., Kim, J.J.: Effect Sizes for Research. Lawrence Erlbaum Associates, Mahwah (2005)

32. Cohen, J.: Statistical Power Analysis for the Behavioral Sciences, 2nd edn. Lawrence Erlbaum Associates, Hillsdale (1988)
33. Kulik, J.A.: Meta-analytic studies of findings on computer-based instruction. In: Baker, E.L., O'Neil Jr., H.F. (eds.) Technology Assessment in Education and Training, pp. 9–33. Erlbaum, Hillsdale (1994)
34. Vinsonhaler, J.F., Bass, R.K.: A summary of ten major studies on CAI drill and practice. Educ. Technol. **12**, 29–32 (1972)
35. VanLehn, K.: The relative effectiveness of human tutoring, intelligent tutoring systems, and other tutoring systems. Educ. Psychol. **46**, 197–221 (2011). https://doi.org/10.1080/00461520.2011.611369
36. Kulik, J.A., Fletcher, J.D.: Effectiveness of intelligent tutoring systems: a meta-analytic review. Rev. Educ. Res. **86**, 42–78 (2016)
37. Fletcher, J.D.: DARPA Education dominance program: April 2010 and November 2010 digital tutor assessments. (IDA Document NS D-4260). Institute for Defense Analyses, Alexandria, VA (2011)
38. Fletcher, J.D.: The value of digital tutoring and accelerated expertise for military veterans. Educ. Technol. Educ. Dev. **65**, 679–698 (2017)
39. Fletcher, J.D., Morrison, J.E.: Accelerating development of expertise: a digital tutor for Navy technical training (IDA Technical report D-5358). Institute for Defense Analyses: Alexandria, VA (2014)
40. Csikszentmihalyi, M.: The Psychology of Optimal Experience. Harper Perennial, London (1990)
41. Salary.com. Network administrator II (2014). http://swz.salary.com/SalaryWizard/Network-Administrator-II-Salary-Details.aspx
42. Ericcson, K.A., Charness, N., Feltovich, P.J., Hoffman, R.R. (eds.): The Cambridge Handbook of Experts and Expert Performance. Cambridge University Press, New York (2006)
43. Hoffman, R.R., Ward, P., Feltovich, P.J., DiBello, L., Fiore, S.M., Andrews, D.H.: Accelerated Expertise: Training for High Proficiency in a Complex World. Psychology Press, New York (2014)
44. Gott, S.P., Lesgold, A.M.: Competence in the workplace: how cognitive performance models and situated instruction can accelerate skill acquisition. In: Glaser, R. (ed.) Advances in Instructional Psychology, pp. 239–327. Erlbaum, Hillsdale (2000)
45. Gott, S.P., Lesgold, A.M., Kane, R.S.: Tutoring for transfer of technical competence. In: Wilson, B.G. (ed.) Constructivist Learning Environments: Case Studies in Instructional Design, pp. 33–48. Educational Technology, Englewood Cliffs (1996)
46. Lesgold, A.S., Lajoie, M., Bunzo, E.G.: SHERLOCK: a coached practice environment for an electronics troubleshooting job. In: Larkin, J., Chabay, R., Scheftic, C. (eds.) Computer Assisted Instruction and Intelligent Tutoring Systems: Establishing Communication and Collaboration, pp. 201–238. Lawrence Erlbaum Associates, Hillsdale (1988)
47. Wetzel-Smith, S.K., Wulfeck, W.H.: Training incredibly complex tasks. In: Cohn, J.V., O'Connor, P.E. (eds.) Performance Enhancement in High-Risk Environments, pp. 74–89. Praeger, Westport (2010)
48. Chatham, R.E., Braddock, J.V.: Training superiority and training surprise. United States Department of Defense, Defense Science Board, Washington, DC (2001)
49. Cohn, J., Fletcher, J.D.: 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 Annual Conference. National Training and Simulation Association, Arlington, VA (2010)

50. Atkinson, R.C., Fletcher, J.D.: Teaching children to read with a computer. Read. Teach. **25**, 319–327 (1972)
51. Suppes, P.C.: The place of theory in educational research. Educ. Res. **3**, 3–10 (1974)
52. Suppes, P., Morningstar, M.: Computer-Assisted Instruction at Stanford 1966-68: Data, Models, and Evaluation of the Arithmetic Programs. Academic Press, New York (1972)
53. Jamison, D.T., Suppes, P., Wells, S.: The effectiveness of alternative instructional media: a survey. Rev. Educ. Res. **44**, 1–67 (1974)
54. Niemiec, R., Walberg, H.J.: Comparative effects of computer-assisted instruction: a synthesis of reviews. J. Educ. Comput. Res. **3**, 19–37 (1987)
55. Atkinson, R.C., Paulson, J.A.: An approach to the psychology of instruction. Psychol. Bull. **78**, 49–61 (1972)
56. Chant, V.G., Atkinson, R.C.: Application of learning models and optimization theory to problems of instruction. In: Estes, W.K. (ed.) Handbook of Learning and Cognitive Processes, vol. 5. Erlbaum Associates, Hillsdale (1978)
57. Sottilare, R.A., Brawner, K.W., Goldberg, B.S., Holden, H.K.: The generalized intelligent framework for tutoring (GIFT). Concept Paper Released with GIFT Software Documentation. U.S. Army Research Laboratory – Human Research & Engineering Directorate (ARL-HRED), Orlando FL (2012)
58. Suppes, P.C.: Modern learning theory and the elementary-school curriculum. Am. Educ. Res. J. **1**, 79–93 (1964)
59. Graesser, A.C., Person, N.K., Magliano, J.P.: Collaborative dialogue patterns in naturalistic one-on-one tutoring. Appl. Cogn. Psychol. **9**, 495–522 (1995)
60. Koedinger, K.R., Corbett, A.T., Perfetti, C.: The knowledge-learning-instruction framework: bridging the science-practice chasm to enhance robust student learning. Cogn. Sci. **36**, 757–798 (2012)
61. Person, N.K., Bautista, L., Graesser, A.C., Mathews, E.C., The Tutoring Research Group: Evaluating student learning gains in two versions of AutoTutor. In: Moore, J.D., Redfield, C. L., Johnson, W.L. (eds.) Artificial Intelligence in Education: AI-ED in the Wired and Wireless Future, pp. 286–293. IOS Press, Amsterdam (2001)
62. Nye, B.D., Pavik Jr., P.I., Windsor, A., Olney, A., Hajeer, M., Hu, X.: SKOPE-IT (shareable knowledge objects as portable intelligent tutors): overlaying natural language tutoring on an adaptive learning system for mathematics. Int. J. STEM Educ. (2018). https://doi.org/10.1186/s40594-018-0109-4
63. Healy, A.F., Kole, J.A., Bourne, L.E.: Training principles to advance expertise. Front. Psychol. **5**, 1–4 (2014)
64. Moreno, R., Mayer, R.: Interactive multimodal learning environments. Educ. Psychol. Rev. **19**, 309–326 (2007). https://doi.org/10.1007/s10648-007-9047-2
65. Fletcher, J.D., Sottilare, R.A.: Shared mental models in support of adaptive instruction for teams using GIFT. Int. J. Artif. Intell. Educ. (2017). https://doi.org/10.1007/s40593-017-0147-y
66. Jones, M.B.: Regressing group on individual effectiveness. Organ. Behav. Hum. Perform. **11**, 426–451 (1974)
67. Sottilare, R.A., Ragusa, C., Hoffman, M., Goldberg, B.: Characterizing an adaptive tutoring learning effect chain for individual and team tutoring. In: Proceedings of the Interservice/Industry Training Simulation and Education Conference, Orlando, FL (2013)