Tech., Inst., Cognition and Learning, Vol. 11, pp. 51–72
Reprints available directly from the publisher
Photocopying permitted by license only

Adapting Learning with Digital Tutoring

J. D. FLETCHER

Institute for Defense Analyses, Science and Technology Division, Institute for Defense Analyses, 4850 Mark Center Drive, Alexandria, VA 22311, USA, e-mail: Fletcher@ida.org

Computer technology has been used for over 50 years to tailor learning experiences to the needs and interests of individual learners at all levels of instruction. It provides adaptation and individualization that is difficult, if not impossible to apply in a classroom of 20–30 students. This article provides a brief background and discussion about adapting instruction for individuals, the use of computer technology to deliver adaptive instruction, and finally the design, development, and assessment of a computer-based tutor developed by the Defense Advanced Research Projects Agency (DARPA) using machine intelligence to capture the benefits of one-on-one tutoring for individuals. This tutor provided novice US Navy sailors with 16 weeks of instruction in information technology. The Tutor accelerated their development to expert-level capabilities and enabled them to outperform, with effect sizes in excess of three standard deviations, other novice learners who received 35 weeks of classroom instruction and sailors with an of average 9 years of experience.

Keywords: Adaptive instruction, Individualized instruction, Digital tutoring, Effectiveness of digital tutoring, Intelligent tutoring systems , Acceleration of expertise, Information systems technology

1 INTRODUCTION

This article concerns the use of computer technology to provide learning experiences – a topic not unfamiliar to readers of this journal. More specifically, it provides an example of adapting learning with a particular approach – using

*Corresponding author: Fletcher@ida.org

computer technology to capture the considerable instructional benefits of one-on-one tutoring. It provides a brief background and discussion about adapting presentation of instruction in general, then the use of computers to adapt instruction to individual learners, and finally the design, development, and assessment of a particular computer-based tutor developed by the Defense Advanced Research Projects Agency (DARPA), which produced learning equivalent to one-on-one human tutoring and a possible breakthrough in the history and application of machine intelligence to provide adaptive education and training.

2 ADAPTIVE LEARNING

Today, and according to the US Department of Education (USDoEd, 2010), adaptive learning clusters around one of three approaches, each suggesting slightly different adaptations: **differentiation**, which adapts the instructional approach, but not its objectives, to groups of learners; **personalization**, which adapts objectives and topics to learners' preferences, and **individualization**, which adapts instruction to learners' abilities, prior learning, and learning progress, but not its objectives. These approaches overlap to some degree, but each provides a somewhat different base for designing and developing instruction.

Choosing which of these approaches to apply may depend on the overall purpose of the instruction. 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 objectives may be a better approach for trainers while personalization with its adjustment of objectives to the learner may be a better choice for educators. Differentiation seems caught somewhere between these two.

This article focuses on individualization. However, and as Table 1 suggests, the differences between training and education are neither rigid nor absolute. Their differences may be characterized along a continuum that we might identify as instruction. Most training includes elements of education and most education includes elements of training. Both physicists and electronic technicians must learn Ohm's Law and algebra, but both may 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. Differences of emphasis and objectives determine where on the continuum of training to education a specific instructional activity belongs and thereby what adaptive strategy to emphasize in designing learning experiences. Despite its need in training, individualization, with its focus on specific objectives, finds a significant role in both.

TABLE 1
Comparison of Education and Training

Education	**Training**
Life Objectives	Job Objectives
Negotiable Objectives	Fixed Objectives
Cost-Effectiveness	Return on Investment
An End in itself	Means to an end
Includes training	Includes education

Most trainers have experience with education and are familiar with the Table 1 issues, but educators may wonder why the issue of training comes up at all. Its presence, relevance, and importance may be defended by the continuum suggested by the table. Both education and training are fundamentally concerned with learning and cognition, and they have much to learn from each other. Although many findings from digital tutoring, here and elsewhere, come from instruction closer to the training than the education end of the continuum, many techniques and findings from training research are as relevant to education as they are to training, particularly as they come together in the daily practice of teaching as they do for boatswains and physicists. These comments are simply to suggest that research findings from either end of the training to education continuum are frequently relevant and applicable to the other. Given this context, we may turn to individualization more directly, its emphasis in this article, and its role in the use of computers in instruction.

3 INDIVIDUALIZED INSTRUCTION

In 1890, William James 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 individuals differ, as they invariably do, then it is likely that their perceptions, learning, and cognition differ. This observation leads to Thorndike's (1906) assertion 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).[1]

[1] This is a remarkably constructivist statement coming from a scholar who is considered to be a resolute stimulus-response psychologist.

These views continue to be supported by empirical research indicating the extent of individual differences that teachers and instructors must deal with in classrooms. Marching through the decades, we find Suppes, Fletcher, & Zanotti (1975, 1976) reporting 1:4 as the ratio in time needed by fastest to slowest learners to achieve elementary school mathematics objectives. Gettinger (1984) 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 held back from valuable learning opportunities and the advanced competencies they could produce while others, equally deserving of learning, struggle to keep up.[2]

Another primary and often noted source of differences in time to learn is prior learning (e.g., Tobias, 2003). Because of the increased variety in prior knowledge acquired by individuals through time and life experience, it is likely that these ratios increase with the age and varied experiences of individuals, as found in higher education and workforce training. In any case, adapting learning to individual differences appears to be a continuing imperative and a particularly difficult issue for classroom education and training at all levels.

The problem of individual differences in background, temperament, ability, and prior knowledge can be eased by classroom practices and heroic efforts of classroom teachers, but only partially. Despite its obvious economic advantages, classroom instruction and its difficulty in attending to individual learners, presents an unavoidable impediment to efficiency and effectiveness in both training and education. Bloom (1984) and his students' research indicated a learning increase of two standard deviations in using tutoring (one instructor working with one learner) rather than classroom instruction – a difference that (roughly and on average) would increase 50th percentile learners to the 98th percentile.

Discussion about Bloom's empirical findings continues, but subsequent research supports the substantial superiority of individual tutoring over classroom instruction in providing adaptive learning (e.g., Cohen, Kulik, & Kulik, C-L, 1982; Evans & Michael, 2006; Graesser, D'Mello, & Cade, 2011; Graesser, Person, & Magliano, 1995). Why then do we not provide an Aristotle for every Alexander and a Mark Hopkins for the rest of us? The answer is obvious. 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 universal provision of individualized, on-demand,

---

[2] Gettinger emphasized, in accord with Carroll (1970), that this difference is not solely due to native ability, but, more precisely to what they both described as individual differences in learning ability.

tutorial instruction. Full natural language, with its use of metaphors, similes, slang, and other peculiarities, may remain beyond the capabilities of computers for some time, but an appreciable range of highly adaptable tutorial dialogue now appears affordably within our reach, if not fully in our grasp.

This possibility suggests a vision of personal computer-based devices (e.g., laptops, telephones) providing individualized instruction, performance aiding, and decision support as tutorial dialogues at anytime and practically anywhere. Also, and aside from algorithms for tutoring and private information about the learner, the necessary (and up to date) subject matter need not be stored locally. It can be collected as required and/or on demand from the global information grid and tailored to the background, needs, evolving capabilities, and even interests of the individual learner (Fletcher, Tobias, & Wisher, 2007).

Like many innovations (e.g., horseless carriages, wireless telegraph), computer-assisted instruction (CAI) began by layering an existing technology (text book based programmed learning) onto another (computers) to provide interactive instruction. Programmed learning applied in CAI is based on processes and frames such as that shown in Figure 1. Typically it uses Keller's Personalized System of Instruction (1968) to determine which set of programmed learning frames a learner should receive. Then, in accord with Crowder's (1959) Intrinsic Programming – as opposed to Skinner's (1954) Extrinsic Programming, it presents instructional frames consisting of text and frames, such as the one illustrated in the figure, to determine the next steps for instruction depending on the learner's response.

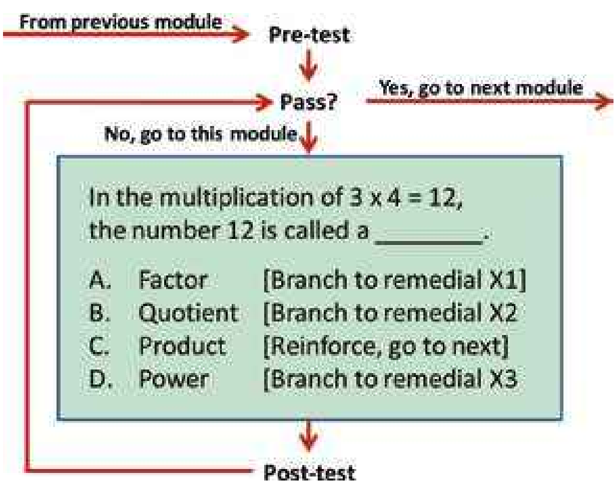

FIGURE 1
Typical Intrinsic Programming Frame.

It is relatively inexpensive to write computer code for these frames, and the approach is still widely used today. Reviews found it to be modestly superior to classroom learning generally improving learning by 50th percentile learners (roughly and on average) to the 66th percentile (e.g., Kulik, Cohen, & Ebeling, 1980; Kulik, C-L, Schwalb, & Kulik, J., 1982). However, frame oriented instruction requires considerable human effort (and expense) to compose because the frames need to fully anticipate and prepare in advance for every likely state of the learner and the instructional system, an effort that is expensive in its use of human time. It was also early found to be impossible – even for something as rudimentary as 2nd grade subtraction (Barr & Feigenbaum, 1982). A more adaptive approach was evidently needed – an approach in which the state of the learner and the instruction would be dynamically modeled, generated, and provided by computer – in real time and as needed for tutorial instruction. This possibility was a primary motivation for the Department of Defense to fund research and development of intelligent tutoring systems or, in terms used here, digital tutors (Fletcher & Rockway, 1986).

4 DIGITAL TUTORING

With support from the Office of Naval Research in the mid-1960's, Wallace Feurzeig 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 (Feurzeig, 1969). Based on initial development of the Mentor language, semantic networks (Quillian, 1969), and his own work on SCHOLAR, Carbonell (1970) identified two basic features that distinguish digital tutors from standard CAI:

- Dynamic information structures in place of pre-programmed, frame-oriented exercises. Information structures, such as those based on ontologies, concept maps, natural-language understanding, and one-on-one tutorial strategies, relieve developers from the need to anticipate every state that might exist for individual learners 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. Because of their generative capabilities, digital tutors allow either the computer or the learner to initiate inquiries during an instructional dialogue. Either the student or the computer can take the initiative in asking questions and posing problems. Tutorial responses by the

computer can be generated and tailored to what the student learned as well as the context within which the inputs occurred. For that matter, the computer could provide guidance and assistance to the learner, even and if necessary, before the learner knows what questions to ask.

These distinctions were based on what Carbonell (1970) called Information Structure Oriented Instruction as opposed to the Ad-Hoc Frame Oriented instruction that used programmed learning techniques. Further, and in contrast to programmed learning approaches, Mentor used mixed initiative dialogues that allowed either the computer or the learner to initiate questions and lines of inquiry. This approach was later the basis for developing, among others, the SOPHIE system to train electronic technicians as illustrated in Figure 2 (Brown, Burton, & DeKleer, 1982). Comments in the figure show a dialogue obtained from a learner troubleshooting an electronic power supply using 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 often pointed out from the 1970s on, digital tutors typically apply three explicit models: (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 of the subject matter and acquisition of the targeted 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 (Fletcher, 1975). Later (4) a model for communication between the learner and the instructional system was found needed (Anderson, Boyle, Corbett, & Lewis, 1990; Gott, Lesgold, & Kane, 1996). 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 evolves dynamically with the student and must be developed and revised in real time, preferably in a stealthy manner (e.g., Shute & Ventura, 2013) with a minimum of explicit testing. It provides the foundation for adapting learning.

5 EFFECTIVENESS OF DIGITAL TUTORING

Assessments of digital tutors generally report greater statistical effectiveness compared to other instructional approaches. However among them, and in other

	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 the prior question.
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. Additionally it 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?	

[a] Adapted from Brown, Burton, & DeKleer (1982)

FIGURE 2
Sample Tutorial Dialogue with Sophie[a]

research as well, there are numerous findings that are statistically significant, but of minor instructional value and/or return on investment time and cost. For that reason, reviews and meta-analyses 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 (e.g., Ellis, 2010; Grissom & Kim, 2005). These differences are notable, and worthy of consideration, but they are not as much an unsettled issue as their interpretation, which may be peculiar to the individuals or organizations using them for decision-making. Table 2 is provided here – with assistance from Cohen (1988), the DoEd What Works Clearinghouse (2010), and

TABLE 2
Overview of Effect Size

Effect Size	Suggested Designation[a]	50^{th} Percentile (Roughly) Raised To …
ES < 0.25	Negligible[b]	60^{th} percentile
0.25 < ES < 0.40	Small	60^{th}–66^{th} percentile
0.40 < ES < 0.60	Moderate	66^{th}–73^{rd} percentile
0.60 < ES < 0.80	Large	73^{rd}–79^{th} percentile
ES > 1.00	Very Large	80^{th} percentile and up
ES > 2.00	Bloom's challenge[c]	98^{th} percentile and up

[a] Extended from suggestions by Cohen (1988).
[b] What Works Clearinghouse (2010).
[c] Bloom (1984).

Bloom (1984) – as a suggested guide for interpreting effect sizes, or at least those reported here.

Evidence of digital tutoring effectiveness is growing. Overall it has been found to be superior to both classroom instruction and other CAI systems. Earlier evaluations of frame-based CAI found that it produces more learning than classroom learning over a variety of subject matter in the area of 0.33 standard deviations (Kulik, 1994). This result would be viewed as a small, but appreciable improvement, given the guidelines in Table 2. More recent meta-analysis of digital tutoring was provided by Fletcher and VanLehn (2011) and Kulik and Fletcher (2016).

VanLehn (2011) reviewed 27 studies of digital tutoring and found that they averaged effect sizes of 0.59, roughly suggesting an appreciable improvement of 50th percentile learners to the 72nd percentile. However, he then investigated improvements based on the precision of the tutoring provided by comparing sub-step, specifically directed tutoring to full-step, more generally directed tutoring. He found an average effect size of 0.40 for sub-step-based tutoring compared to an 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 but improve (roughly and on average) to the 78th percentile under more general, less specific tutorial interactions.

Additional research may better account for this result, which could have been due to the need for students to consider, i.e., reflect, more carefully under step-based tutoring than under sub-step-based tutoring how to solve the problem before them and transfer what they have learned to similar problems. In either case it seems reasonable to expect digital tutoring to produce practical effects well in excess of those found in classroom instruction or in frame-based CAI.

A larger and more recent analysis by Kulik and Fletcher (2016) found an average effect size of 0.66 for 50 digital tutors, with data ranging from -0.34 to 3.18 (after winsorizing for outliers) – a finding 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 applications of paper-based programmed learning techniques in CAI.

6 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 language and program. However, DARPA's mission is to develop high payoff research that is too risky and expensive to be developed by the military Service laboratories. An example is 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 the acquisition of expertise, well beyond novice or journeyman levels, by learners beginning 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 systematic empirical testing. It was not focused on verifying any particular theory of learning, cognition, and/or instruction. The developers had spent a number of years studying one-on-one tutoring by humans. Findings from that work provided initial approaches in designing the DARPA Tutor.

The funding provided by DARPA allowed more extensive assessment of these approaches using IT novices of about the same age, education level, and mental capabilities as Sailors who were recently recruited and assigned to a Navy IT school. The Tutor was then developed lesson by lesson in sessions with detailed video and human recordings to determine what worked and what did not. These sessions were repeated as necessary to develop instructional activities and interactions that reliably produced the intended learning objectives. This iterative approach established a problem-based learning environment using generative information structures functionally similar to those of human tutors who have expertise in the subject matter and one-on-one instruction. Instruction was organized around specific topics in IT problem solving.

Experts in these topics were identified and vetted based on their success, publications, and reputation.[3] These subject matter experts were auditioned in 30-minute sessions in whatever subject matter they chose to tutor learners who were representative of new Navy sailors. The intention was to base (or "clone") the Digital Tutor on the practices of individuals who were expert in both an IT topic and in one-on-one tutoring.

These sessions were used to select 24 tutors in specific IT topics who then trained 15 IT qualified sailors, 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. Initial design and development of the Digital Tutor was then based on the tutorial sessions with these sailors. These sessions, which were extensively reviewed and assessed, served as the basis for tutorial instruction provided by the Digital Tutor.

Based on this work, 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 or give the correct answer or 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 are or are not relevant to the current problem; and
- Require logical, causal, and/or goal-oriented reasoning in reviewing or querying correct and incorrect actions taken by the learner to solve problems.

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 human-tutored and 17 classroom instructed sailors. The tutored sailors averaged

3 This work was performed in Silicon Valley.

77.7 points compared to 39.7 points for the classroom training sailors on this test, which indicated an effect size of 2.48 in their favor (Fletcher & Morrison, 2014).

Operationally, the design of the Digital Tutor, was based on these tutorials. It emphasizes:

- Active, constant interaction with learners – which fostered the “flow” that is found in computer-based games (Csikszentmihalyi, 1990);
- Capture in digital form of the processes and practices of one-on-one tutoring;
- Problem solving in authentic environments – leaners used actual Navy systems, not simulations of these systems[4], also the 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 Tutor uses information structures to:

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

Mirroring its development strategy, the Tutor’s instructional approach is spiral. It presents conceptual material that is immediately followed and applied in solving problems intended to be comprehensive and authentic. Learners interact directly with IT systems 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.

[4] These systems could, if necessary, be quickly repaired or re-built by the Tutor.

  – 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.

7 EFFECTIVENESS OF THE DARPA DIGITAL TUTOR

The evolving Digital Tutor was evaluated twice, after 4 and then 8 weeks of Tutor development, with different classes of 12 novice sailors (Fletcher & Morrison, 2014). A daily schedule consisted of 6 hours 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.

The first assessment 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 IT training and with that of 10 Navy IT instructors. This study 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 course and an effect size of 1.32 in their favor compared to their instructors.

The next assessment compared 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 1.31 over their instructors. All these differences were statistically significant ($p< 0.05$).

These differences aside, the presence of experienced Navy ITs was essential for this training. They resolved difficulties in human-computer communication, managed the study halls, and, especially provided examples of Navy bearing and cultural for the younger sailors. "Sea stories" might be viewed as little more than embellishment, 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 civilians for military service. In subsequent years, the differences in learner and instructor subject matter expertise might be readily overcome by requiring Navy ITs newly assigned for shore duty as instructors at the school to complete the 16-week digital tutor as a first order of business.

A final assessment was performed after another representative group of 12 sailors had completed training with the final 16-week version of the Tutor (Fletcher & Morrison, 2014). The DARPA challenge was to produce in 16 weeks (the usual time for ab initio IT training) novice sailors who were superior in skill and knowledge to (a) other sailors trained using conventional classroom and laboratory practice, and (b) ITs with years of experience in the Fleet.

As it turned out, the final assessment involved new sailors trained using the DARPA Digital Tutor, other new sailors trained by 35 weeks using the Navy's classroom based Information Technology Training Continuum (ITTC), and ITs with an average of 9.2 years of Fleet experience. Again, sailors who had just finished recruit training were assigned at random to each of the two training groups (DT and ITTC 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 hours of IT trouble shooting skill testing, and 4 hours 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. A full description of this testing was provided by Fletcher & Morrison (2014).

IT troubleshooting in response to trouble tickets was the most important component of the training to the Navy in preparing these novice sailors for their Navy IT occupation. It was intended to resemble Fleet IT requirements as closely as possible. Results of the Troubleshooting testing are shown in Figure 3.

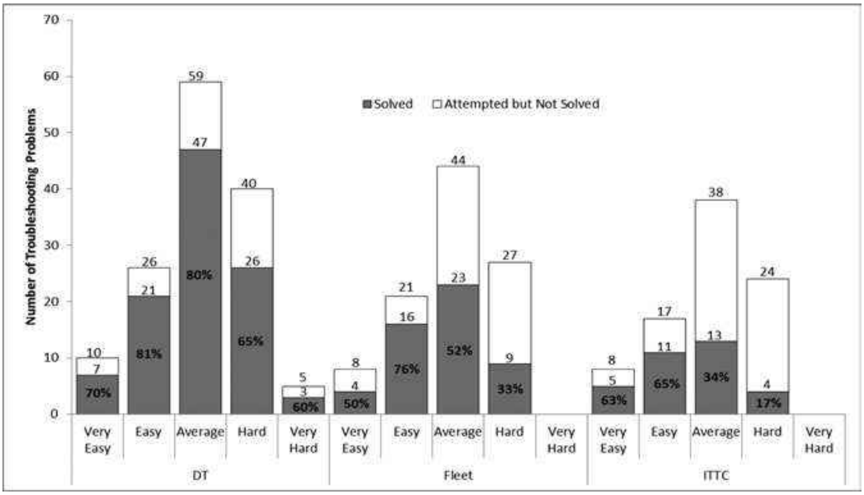

FIGURE 3
Troubleshooting problems solved by DT, Fleet, and ITTC teams

In both cases, the effect sizes are unusually large. Troubleshooting capability was the main focus of the assessment because the instruction was focused on what the sailors could do. In additional analyses, knowledge was 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.

The effect sizes are unusually large, but they have stood up to considerable scrutiny. Additional or other testing might produce findings of different magnitude, but they are still likely to be sizable. Considerably more description of this assessment, additional data and findings are provided by Fletcher and Morrison (2014) as well as in a forthcoming publication.

An approximate replication was provided by an assessment of an 18 week version of the Tutor used to train 100 military veterans (Fletcher, 2017). As Table 3 shows, most of the veterans were unemployed before taking this course. There were no academic dropouts from the course, which was completed by 97 veterans of the veterans.[5] All 77 of the graduates who sought employment were hired with an average annual salary of $73,000, which is equivalent to civilian employment reserved for IT technicians with 3–5 years of IT experience (Salary.com, 2014).

[5] Two veterans dropped out because of health reasons. One dropped out because of a death in the family.

TABLE 3
Characteristics of 101 Veterans Accepted for Training [a]

Average Years of Separation from Service	**5.20**
Avg Age	30.5
Married	30
Armed Forces Qualiification Test	87.1
Full Time Employment	11
Part Time Employment	45
Prior Civilain IT Instruction	8
High Schoo/GED Degree	45
AA Degree	11
BA/BS	44
Other	1
Prior Military IT Instruction	4

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

Results from return on investment analysis are shown in Figure 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

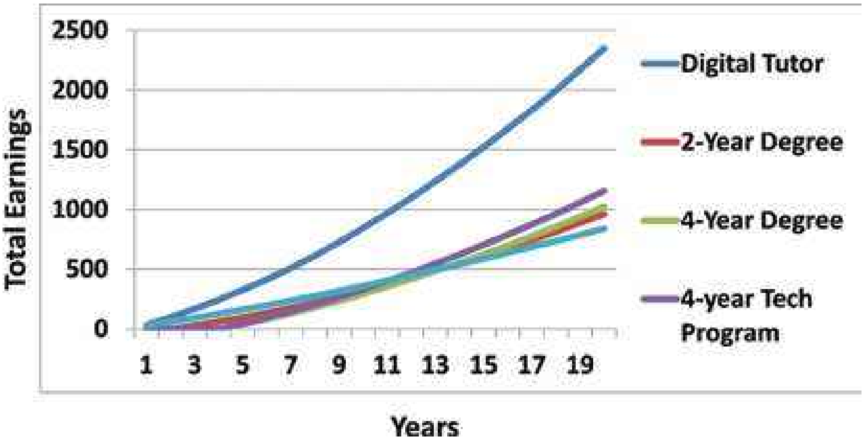

FIGURE 4
Monetary Return to US Government Per Individual from Support Provided for Education and/or Training ($000)

the 18-week digital tutoring program than government support for a either a 2 or 4 year degree and even for a program that receives no government support.

8 DISCUSSION

The term ‘adaptive learning’ raises a question of what are we adapting to -- a learner’s interests, ability, learning progress, prior learning, age, objectives (personal and instructional), temperament, Piagetian stage, or something else? All of these may be reflected and applied in the dynamic models of the learner generated by tutors, human and digital.

One-on-on tutoring may therefore be the ultimate in adapting instruction/ learning to individual learners. As noted, such tutoring by humans is prohibitively expensive for most education and training. However, computers may put this tutoring within economic reach. In some cases, such as the early (and still continuing) uses of programmed learning in the form of frame-based, intrinsic programming, their use may produce marginal learning improvements for learning a variety of instructional objectives (Kulik, C.-L. & Kulik, J.A., 1991). In other cases, such as the use of drill and practice techniques to achieve early, ab inito, learning objectives, they may be both effective and cost-effective (Fletcher, Hawley, & Piele, 1990). However, as reviews of properly identified digital tutors have found, tutoring appears to be significantly more effective in achieving the deeper and more substantive objectives that are required for retention and transfer of learning. For that matter, assessment of the DARPA Digital Tutor suggests that they are or will soon become as effective as human one-on-one tutoring – if not more so – and thereby affordable providers of adaptive, individualized learning.

Digital tutoring continues to be more expensive to produce than drill and practice or programmed learning approaches, but, by relieving the need for extensive on-the-job training, the return on investment, particularly in reducing IT problems and eliminating the need for on-job-training was found by Cohn and Fletcher (2010) to be substantial.

Recent research is developing techniques to reduce the time and cost to developed these systems. Particularly notable is work being performed cooperatively by the Institute for Intelligent Systems at the University of Memphis and the Army Training Research Laboratory’s Human Research and Engineering Directorate to develop the Generalized Intelligence Framework for Tutoring (GIFT) (Sottilare, Brawner, Goldberg, & Holden, 2012). GIFT is described elsewhere in this issue and recommended for those interested in the “authoring” of digital tutoring. Inevitably, these development costs will inevitably decrease if and as more tutors are built along these same lines.

No "magic sauce" or specific academic theory was used to produce the 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. In doing so, its development was in the mode of Herbert Simon's "Sciences of the Artificial" (1996), which recommends extensive trial and error experience to develop theory, in contrast to the common academic practice of beginning early on with theory and attempting to validate it.

Certainly theory for instruction is essential (e.g., Suppes, 1974), 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 Tutor uses a problem-solving approach – expository text followed by problem solving. Characteristically, solution paths for the problems were devised and carried out by individual learners. They were not based on solutions prescribed by experts. Because the Tutor was intended to produce expert-level capabilities and because different experts may follow different solution paths to solve problems, sufficient capabilities were built into the Tutor to allow it to correct, amend, and/or adopt whatever solution path the learner was using to solve the problem.

Table 4 reports an average effect size of 0.08 for using digital tutoring to present rudimentary material, such as nomenclature, basic facts, and introductory procedures, which seems better suited to drill and practice approaches. Material of this sort must be learned, often by rote, to give beginners a foothold in learning any new subject matter. The finding of such a small effect size is in contrast to the average effect size around 0.75 for post-secondary content which, although starting from rudiments, rapidly ascends to more abstract, conceptual levels of learning, which require more frequent and substantive tutorial interactions between learner and instructor.

These findings are illustrated in Figure 5, which uses categories adapted from Anderson and Krathwohl's (2001) two-dimension expansion of Bloom's 1984 hierarchy of learning.[6] The figure suggests first that there is in almost all learning

[6] Both Anderson and Krathwohl worked on Bloom's original taxonomy of educational objectives.

TABLE 4
Effect Sizes for Four ITS Systems Assessed for Conceptual and Rudimentary Learning

Source	Concepts	Rudiments
Graesser, Moreno, et al. (2003)	0.34	0.00
Koedinger et al. (1997)	0.99	0.36
Person et al. (2001)	0.30	0.03
VanLehn et al. (2005)	0.95	-0.08
Average (Standard Deviation)	0.65 (0.326)	0.08 (0.168)

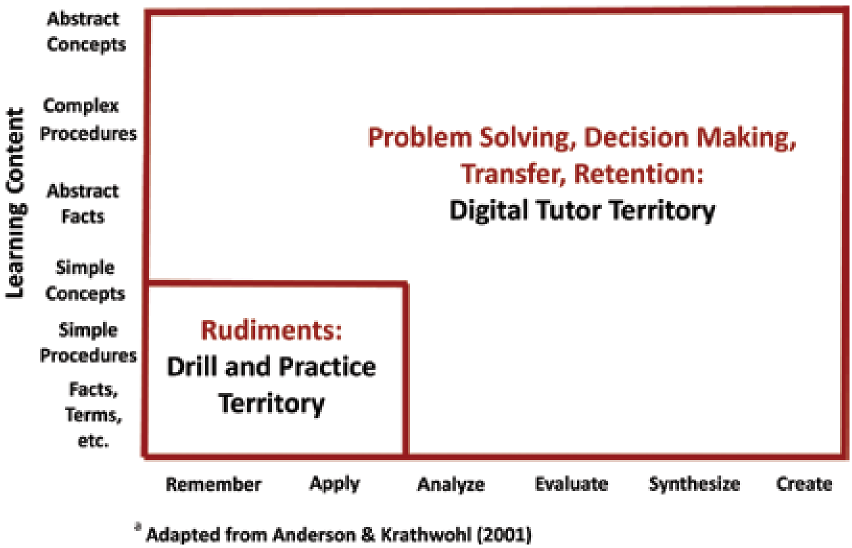

FIGURE 5
Suggested Roles for Drill and Practice and Digital Tutoring.[a]

an essential role for drill and practice and, second, that digital tutoring finds its value, if not its necessity, in applying these rudiments to the more conceptual and abstract levels of learning needed for transfer and retention as reported in many studies (e.g. Gott, Lesgold, & Kane, 1996; Healy, Kole, and Bourne, 2014).

In sum, rudiments may be better (and more economically) taught using drill and practice techniques and instruction involving abstract and conceptual learning may be best provided by digital tutors. As for the heretical and unpopular view that drill and practice has its place – wide experience, including that from the gaming world where the development of “flow” (Csikszentmihalyi, 1990) among individuals of every age indicates that they enjoy playing against themselves (or others) and seeing improvement. “Drill and kill” is a clever deprecation of the term, but it does not characterize computer-based drill and practice done well.

Kulik and Fletcher (2016) report four instances where an ITS was assessed twice, once focused on rudimentary learning (basic facts, nomenclature, and simple procedures) and another focused on deeper, conceptual learning. The findings in Table 4 suggest that ITS systems may be better suited to providing deep rather than rudimentary learning, which, as Figure 5 suggests, may be better relegated to the drill and practice activities covering the nomenclature, introductory procedures, and simple concepts needed to prepare new learners to any new course of study. Given this background using relatively simple (and inexpensive) instructional approaches learners are then ready to pursue the deeper, more conceptual understanding which is needed for retention of the material, expansion of problem solving ability, and transfer to related problems and areas of the subject matter. The DARPA Digital Tutor was designed and developed with these instructional approaches in mind.

In summary, 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. How best to move it from the laboratory into the field, however, remains undetermined.

REFERENCES

Anderson, J. R., Boyle, C. F, Corbett, A., & Lewis, M. W. (1990). Cognitive modeling and intelligent tutoring. *Artificial Intelligence, 42,* 7–49.

Anderson, L. W. & Krathwohl, D. R. Eds., (2001), *A Taxonomy for Learning, Teaching, and Assessing: A Revision of Bloom’s Taxonomy of Educational Objectives*, p. 28. Copyright 2001 by Allyn & Bacon.Barr,

Barr, A. & Feigenbaum, E. (1982). Buggy. *Handbook of Artificial Intelligence*, *Volume 2* (pp. 279–284). Los Altos, CA: William Kaufmann, Inc.

Bloom, B. S. (1984). The 2 sigma problem: The search for methods of group instruction as effective as one-to-one tutoring. *Educational Researcher*, *13*(6), 4–16.

Brown, J. S., Burton, R. R., & DeKleer, J. (1982). Pedagogical, Natural Language and Knowledge Engineering in SOPHIE I, II, and III. In D. Sleeman and J. S. Brown (Eds.), *Intelligent Tutoring Systems* (pp. 227–282). New York: Academic Press.

Carbonell, J. R . (1970) AI in CAI: An Artificial Intelligence Approach to Computer-Assisted Instruction. *IEEE Transactions on Man-Machine Systems*, *11*, 190–202.

Carroll, J. B. (1970). Problems of measurement related to the concept of learning for mastery. *Educational Horizons*, *48*, 71–80.

Cohen, J. (1988). *Statistical Power Analysis for the Behavioral Sciences*: *2nd Edition*. Hillsdale, NJ: Lawrence Erlbaum Associates.

Cohen, P. A., Kulik, J. A., & Kulik, C.-L. C. (1982). Educational outcomes of tutoring: A meta-analysis of findings. *American Educational Research Journal*, *19*, 237–248.

Cohn, J., & Fletcher, J.D. (2010). What is a pound of training Worth? Frameworks and practical examples for assessing return on Investment in training. *Proceedings of the InterService/Industry Training, Simulation and Education Annual Conference*. Arlington, VA: National Training and Simulation Association.

Crowder, N. A. (1959). Automatic teaching by means of intrinsic programming. In E. Galanter (Ed.), *Automatic Teaching: The State of the Art* (pp.109–116). New York, NY: John Wiley & Sons.

Csikszentmihalyi, M. (1990). *The Psychology of Optimal Experience*. London: Harper Perennial.

Ellis, P.D. (2010). *The Essential Guide to Effect Sizes*. Cambridge, UK: Cambridge University Press.

Evans, M., & Michael, J. (2006). *One-on-one Tutoring by Humans and Computers.* Mahwah, NJ: Lawrence Erlbaum.

Feurzeig, W. (1969). *Computer Systems for Teaching Complex Concepts* (BBN Report 1742). Cambridge, MA: Bolt Beranek & Newman, Inc. (DTIC AD 684 831)

Fletcher, J.D. (1975). Models of the learner in computer-assisted instruction. *Journal of Computer-Based Instruction*, *3*, 118–126.

Fletcher, J. D. (2011). *DARPA Education Dominance Program: April 2010 and November 2010 Digital Tutor Assessments*. IDA Document NS D-4260. Alexandria, VA: Institute for Defense Analyses.

Fletcher, J.D, (2017). The value of digital tutoring and accelerated expertise for military veterans. *Education Technology Education and Development*, *65*, 679–698.

Fletcher, J.D., Hawley, D.E., & Piele, P.K. (1990). Costs, effects, and utility of microcomputer assisted instruction in the classroom. *American Educational Research Journal*, *27*, 783–806.

Fletcher, J.D., & Morrison, J.E. (2014). *Accelerating Development of Expertise: A Digital Tutor for Navy Technical Training* (D-5358). Alexandria, VA: Institute for Defense Analyses.

Fletcher, J.D., & Rockway, M.R. (1986). Computer-based training in the military. In J.A. Ellis (Ed.) *Military Contributions to Instructional Technology* (pp. 171–222). New York, NY: Praeger Publishers.

Fletcher, J. D., Tobias, S., & Wisher, R. L. (2007) Learning Anytime, Anywhere: Advanced Distributed Learning and the Changing Face of Education. *Educational Researcher*, *36*(2), 96–102.

Gettinger, M. (1984). Individual differences in time needed for learning: A review of literature, *Educational Psychologist*, *19* (1), 15–29, DOI: 10.1080/00461528409529278

Gott, S. P., Lesgold, A., & Kane, R. S. (1996). Tutoring for transfer of technical competence. In B. G. Wilson (Ed.), *Constructivist learning environments: Case studies in instructional design* (pp. 33–48). Englewood Cliffs, NJ: Educational Technology.

Graesser, A. C., D'Mello, S. K. .& Cade, W. (2011). Instruction based on tutoring. In R. E. Mayer & P.A. Alexander (Eds.). *Handbook of Research on Learning and Instruction*, 408–26. New York, NY: Routledge Press.

Graesser, A. C., Moreno, K., Marineau, J., Adcock, A., Olney, A., & Person, N., & The Tutoring Research Group. (2003). AutoTutor improves deep learning of computer literacy: Is it the dialog or the talking head? In U. Hoppe, F. Verdejo, & J. Kay (Eds.), *Proceedings of Artificial Intelligence in Education* (pp. 47–54). Amsterdam: IOS Press.

Graesser, A.C. Person, N.K., & Magliano, J.P. (1995). Collaborative dialogue patterns in naturalistic one-on-one tutoring. *Applied Cognitive Psychology 9*, 495–522.

Grissom, R.J., & Kim, J.J. (2005). *Effect Sizes for Research*. Mahwah, NJ: Lawrence Erlbaum Associates.

Healy, A. F., Kole, J. A., & Bourne L.E. (2014). Training principles to advance expertise. *Frontiers in Psychology*, 5, 1–4.

James, W. (1890/1950). *Principles of Psychology: Volume I*. New York: Dover Press.

Keller, F.S. (1968). Goodbye, teacher. … *Journal of Applied Behavior Analysis*, *1*, 79–89.

Koedinger, KR., Anderson, J.R., Hadley, W.H., Mark, M.A. (1997). Intelligent Tutoring Goes To School in the Big City. *International Journal of Artificial Intelligence in Education*, *8*, 30–43.

Kulik, J. A. (1994). Meta-analytic studies of findings on computer-based instruction. In E. L. Baker & H. F. O'Neil, Jr. (Eds.), *Technology assessment in education and training* (pp. 9–33). Hillsdale, NJ: Erlbaum.

Kulik, C.-L. C., & Kulik, J. A. (1991). Effectiveness of computer-based instruction: An updated analysis. *Computers in Human Behavior*, *7*, 75–94.

Kulik, J. A., Cohen, P. A., & Ebeling, B. J. (1980). Effectiveness of programmed instruction in higher education: A meta-analysis of findings. *Educational Evaluation and Policy Analysis*, *2*, 51–64.

Kulik, J.A., & Fletcher, J.D. (2016). Effectiveness of Intelligent Tutoring Systems: A Meta-analytic Review. *Review of Educational Research*, *86*, 42–78.

Kulik, C-L.C., Schwalb, B.J., & Kulik, J.A. (1982). Programmed instruction in secondary education: A meta-analysis of evaluation findings. *Journal of Educational Research*, *75*, 133–138.

Person, N. K., Bautista, L., Graesser, A. C., Mathews, E. C., & The Tutoring Research Group. (2001). Evaluating student learning gains in two versions of AutoTutor. In J. D. Moore, C. L. Redfield, & W. L. Johnson (Eds.), *Artificial intelligence in education: AI-ED in the wired and wireless future* (pp. 286–293). Amsterdam: IOS Press.

Quillian, M.R. (1969). The teachable language comprehender: A simulation program and theory of language. *Communications of the ACM*, *12*(8), 459–75.

Salary.com. (2014). Network Administrator II. Retrieved from http://swz.salary.com/SalaryWizard/Network-Administrator-II-Salary-Details.aspx?

Shute, V. & Ventura, M. (2013). *Stealth Assessment: Measuring and Supporting Learning in Video Games*. Cambridge, MA: MIT Press.

Simon, H. A. (1996). *The sciences of the artificial* (3rd ed.). Cambridge, MA: MIT Press.

Skinner, B. F. (1954). The science of learning and the art of teaching. *Harvard Education Review*, *24*, 86–97.

Sottilare, R.A., Brawner, K.W., Goldberg, B.S. & Holden, H.K. (2012). *The Generalized Intelligent Framework for Tutoring* (*GIFT*). Concept paper released as part of GIFT software documentation. Orlando, FL: U.S. Army Research Laboratory – Human Research & Engineering Directorate (ARL-HRED).

Suppes, P.C. (1974). The place of theory in educational research. *Educational Researcher*, *3*, 3–10.

Suppes, P., Fletcher, J. D., & Zanotti, M. (1975) Performance models of American Indian students on computer-assisted instruction in elementary mathematics. *Instructional Science*, *4*, 303–313.

Suppes, P., Fletcher, J. D., & Zanotti, M. (1976). Models of individual trajectories in computer-assisted instruction for deaf students. *Journal of Educational Psychology*, *68*, 117–127.

Thorndike, E. L. (1906). *Principles of Teaching*. New York, NY: A. G. Seiler & Company.

Tobias, S. (2003). Extending Snow's conceptions of aptitudes. *Contemporary Psychology*, *48*, 277–279.

USDoEd (2010). *Transforming American Education Learning Powered by Technology. National Education Technology Plan 2010*. Washington, DC: Department of Education, Office of Educational Technology.

VanLehn, K. (2011). The relative effectiveness of human tutoring, intelligent tutoring systems, and other tutoring systems. *Educational Psychologist*, *46*, 197–221. doi: 10.1080/00461520.2011.611369

USDoEd (2010). *Transforming American Education: Learning Powered by Technology. National Education Policy Plan*. Washington DC: US Department of Education.