Intelligent Tutoring Systems:
Prospects for Guided Practice and Efficient Learning

H. Chad Lane
Research Scientist

Institute for Creative Technologies
University of Southern California
http://www.ict.usc.edu
lane @ ict.usc.edu

July 27, 2006

Rapid technological advances and compounding complexities of the modern world have profound implications for all levels of education and training in the United States and around the world. Intelligent tutoring systems (ITS) represent an important class of educational technology poised to play particularly critical role helping learners acquire the skills needed to succeed. This paper argues why this is the case, describes existing ITS technologies and functionalities, summarizes current research streams, and highlights underrepresented areas of research that (in the author's opinion) will be essential to the training and education of the Soldiers and leaders of tomorrow.

The new science of learning

Researchers in the learning sciences seek to uncover fundamental principles of human learning. About 30 year's worth of such findings are summarized in the National Research Council's report *How People Learn* (HPL) (Bransford, Brown, & Cocking, 2000). In this section, we briefly highlight two (of many) consensus points highlighted by this team that have a particular relevance to ITS research.

HPL stresses the importance of helping learners develop a *deeper conceptual understanding* when they are learning a new domain. This means going beyond facts and procedures to thinking about applicability conditions and dynamic modification of such knowledge to fit new situations. "When students gain a deeper conceptual understanding, they learn facts and procedures in a much more useful and profound way that transfers to real world settings" (p.2, Sawyer, 2006). Allowing learners to actively participate in their own learning is essential to this aim. This guideline is broadly supported by many theories of learning (most prominently those of pioneers such as Piaget and Vygostky). Another finding highlighted in HPL is the *importance of reflection*, or metacognition, in learning. In addition to acquiring factual and procedural knowledge (with deep conceptual understanding, of course), learners should also be improving their ability to learn. Reflective skills, such as planning, questioning, explaining, and criticizing, are generally highly developed in experts but not novices. To accelerate the maturation to expertise, then, it is essential to create learning environments that encourage and support these kinds of activities. Clearly, active learning and reflection go hand-in-hand and should be high priority considerations in the development of computer-based learning environments.

Human and intelligent tutoring

There are no known forms of education as effective as a professional human tutor. Students working one-on-one with expert human tutors often score 2.0 standard deviations – roughly two grade levels – higher than students in a conventional classroom (Bloom, 1984). In contrast, the very best intelligent tutoring systems achieve learning gains of about 1.0 standard deviations (Anderson et. al., 1995; VanLehn et. al., 2005). The best computer-aided instructional systems – computer tutors that do *not* use techniques of artificial intelligence (AI) – produce learning gains of about .4 standard deviations (Niemiec & Walberg, 1987).

Why is tutoring effective?

Although a precise answer to the question of why tutoring is more effective than other forms of instruction has remained elusive, most explanations focus on the fact that the best tutors balance the need for active participation of the student with the provision of guidance. This means the student does as much of the work as possible while the tutor provides just enough feedback to minimize frustration and confusion (e.g., Merrill et. al., 1992). Also, effective tutoring has less to do with improved didactic explanations on the part of the tutor and more to do with the interaction between the tutor and student. Chi et. al. (2001) conclude that “students’ substantive construction from interaction is important for learning, suggesting that an ITS ought to implement ways to elicit students’ constructive responses” (p. 518). It is a common pattern in ITS research to first identify effective learning events and patterns in human tutoring, then attempt to emulate them in an ITS.

Classification of ITS technologies

One way to organize tutoring systems is around what role they are intended to play. At one end of the spectrum, some systems are intended to replace a textbook or classroom instruction to deliver domain content for the first time to a student (e.g., intelligent hypermedia systems). At the other end are systems designed to directly support *practice* (sometimes described as “homework helpers”). These usually complement an existing instructional component such as lectures. Although very few systems sit on the edge of this spectrum, ITS research tends to lean to the practice end. Indeed, practice is when “the rubber hits the road” in learning: it represents a volatile time when knowledge gaps are revealed and skills are automatized. Modern theories of learning stress the critical role of practice and most highlight the importance of feedback because of the risks of unguided learning (Kirschner et. al., 2006; Clark, 2004).

Tutoring systems typically support practice in one of two ways. *Product* tutors evaluate final outcomes, such as an essay or a mission plan. Typically, a student works on a solution until it is deemed complete, then submits it for feedback. The ITS analyzes the solution by looking for flaws, omissions, or sub-optimal elements. Some advanced product tutors are able to reverse engineer solutions using techniques such as *plan recognition* to identify the reasoning that likely underlies the solution. An inherent weakness of product tutors is that students might become stuck before they are able generate a solution. More interactive and pro-active systems that provide support while the student is working towards a solution are best described as *process* tutors. This is perhaps the most familiar category because most human tutors operate in this mode – the student is observed step-by-step, feedback and hints are given, questions are asked and answered, and so on.

Recurring ITS research areas

A core collection of "good old fashioned" ITS research areas has managed to stay in favor throughout the years. Broad surveys of the field conducted roughly 10 years ago (Youngblut, 1994; Shute & Psotka, 1996) and more recently (Loftin, 2004) highlight several recurring themes, many that date back to the earliest years of ITS research:

- *Learner or student modeling:* General problems in this area include diagnosis of misconceptions, tracking of learning over time, representation of faulty (i.e., "buggy") reasoning, "open" learner modeling, and affective/emotional modeling. Learner models can provide assessments to instructors, used to generate appropriate problems, and be the basis for individualized, adaptive instruction (although this remains an unrealized goal).
- *Natural language dialogue:* Some of the earliest tutoring systems attempted to use techniques of natural language processing to simulate human-human tutorial dialogues. Even with intermittent periods of low activity, this stream of research has not gone away. Recent support from ONR and NSF helped produce some modernized systems shown to enhance learning as a direct result of improved dialogue quality.
- *Cognitive modeling:* Research in this area generally involves creation of plausible symbolic representations of the rules and strategic thinking needed to solve problems in a domain. Resulting models are used to evaluate student actions, generate feedback, and provide a basis for learner modeling (Anderson et. al, 1995). Tutoring itself can be treated as a task, and so researchers have also built cognitive models of expert tutoring.
- *Complete systems and evaluations:* ITS research overlaps significantly with the learning sciences, and so thousands of systems have been built and many hundreds evaluated to answer research questions. This trend should snowball in the coming years resulting in more effective tutors and continued contributions to the science of learning.
- *Authoring tools, knowledge acquisition, and development tools:* The burden of creating an ITS was quickly realized, and thus approaches to reducing ITS development time (e.g., encoding expertise, teaching strategies, and domain models) began to surface in the early 80's. This continues to be a focus area and is discussed in the next section.

This list encompasses many subcategories, but is certainly not complete. Loftin et. al. (2004) include (in addition) learning strategies, system design, and collaborative learning environments on their list of recurrent ITS research topics. These problems are recurring because, in part, they have resisted truly general solutions. Interweaving complexities such as broad ranging domains of interest, varying learning goals and contexts, and learner differences all contribute to this resistance and have led some to question the efficacy of seeking truly general solutions. A better approach may be to seek to build specialized components that address certain classes of educational problems – this is also one of the conclusions reached by Loftin et. al. (2004).

The first 20-30 years of ITS research produced a large body of AI-based approaches to building educational software. Tutoring systems proved their ability to be involved in learning in ways that other educational software could not. For example, detailed cognitive models proved that students could get help with "mental" steps involved in problem solving. The clear stumbling block was the lack of adoption and large scale transition of ITS technology into schools (see Koedinger et. al. (1997) for a rare counterexample). This was in part due to the massive effort and special skills required to build an ITS, which motivates more attention to authoring tools.

Recent trends and developments in ITS research

The last section summarized several traditional areas of ITS research that continue to receive attention from the community. In this section, we unpack the most prominently represented of these research themes and describe several other areas of interest that have emerged recently.

Learner modeling. In the last five years, a great deal of attention has been given to the modeling of the *affective state* of learners. Most of this effort has targeted *motivation* since there is evidence form the learning sciences that (1) expert human tutors *do* manage the motivational and emotional states of learners, and (2) instruction can be adjusted according to motivation in ways that improve learning. Often using the highly detailed measurements a computer environment can provide (e.g., time between keystrokes), researchers have built algorithms that translate these patterns into evidence about affective states. In some cases, motivation has been tied to ability and help-seeking tendencies while in others feedback frequency has been adjusted based on the system's motivational state estimate of the student.

O*pen learner modeling* is an extension of traditional learner modeling that makes the model a visible and interactive part of the learning environment. In other words, the display includes a representation of the system's internal belief of the student's knowledge state. A common visualization used is the progress bar. As a student solves problems in a domain, each action is tracked and treated as positive or negative evidence towards a belief domain elements are understood (or not). The progress bars move in one direction or the other, all for the student to see. It is often argued that this inherently motivating because students are usually given the chance to "challenge" the model, essentially telling the ITS "I want a problem to solve to prove I possess skill X." Open learner models are often argued to support reflection and active learning because in order to challenge the model, students must assess their own understanding and decide how to work through the curriculum.

Authoring systems. As discussed previously, if ITS technology is to make its way out of the lab on a large scale, authoring tools will need to be available for end users who want to build new tutors or tweak the system's behaviors based on what they find in the field. A recent book on state of the art authoring tools (Murray, Blessing, & Ainsworth, 2003) makes it clear that although the many existing systems have been successful, all remain research prototypes. This is not true for authoring systems that focus solely on creating and modifying domain *content*, but rather in the case of authoring tutorial and expert *knowledge*, significant hurdles remain.

A particularly promising approach to authoring is based on the idea of *authoring by example*. The basic idea is that rather than encoding domain expertise and tutoring knowledge in an AI programming language, the author instead *demonstrates* ideal solutions. To create feedback messages, the user specifies what the system should say to a student at various points of the demonstration. To handle mistakes, the author simply labels parts of the demonstration as errors, then again authors appropriate feedback messages. This is the approach taken by the Cognitive Tutoring Authoring Tools (CTAT) project. Preliminary testing has shown an authoring speedup of between 1.4 and 2 times over a "reduced" version of the tool lacking the demonstration capability (Aleven et. al., 2006). Because the demonstrated models tend to be overly rigid, researchers are also exploring the use of machine learning techniques in attempts to infer cognitive models from series of demonstrations.

Group, collaborative, and online learning. The advent of the internet and relative ease of networking computers together has radically advanced the state of the art in collaborative learning. ITS work in this area, including intelligent support for team training, has historically been quite limited (Loftin et. al., 2004), but has seen dramatic increases in the last few years thanks in part to the successes of the Computer Supported Collaborative Learning (CSCL) community. Jermann, Soller, & Muehlenbrock (2001) point out that early CSCL systems were essentially networked work environments that performed "mirroring" of actions so all participants could be aware of actions taken in some community workspace.

Advanced CSCL systems go beyond mirroring to provide deeper supports, such as *rate* data (showing how fast collaborators are at completing tasks), social networking tools (to reveal level of communication between participants, for example), and problem solving monitoring with feedback (to offer guidance to individual team members). Because robust free-form natural language understanding is an unsolved problem, advice-giving CSCL systems tend to use other techniques to monitor communicative activities. One approach is to require the use of *sentence openers* (such as "I agree, but…" or "Do you know…") which can provide a deep enough level of intentional information to track collaboration patterns.

Evaluations of CSCL systems have shown that many of the expected benefits, such as increased motivation and participation, have not been realized. Studies repeatedly reveal problems such as low participation and communication rates, satisfaction, and limited learning. To deal with some of these problems, ITS researchers have recently focused on a variety of approaches, such as improving team visualization, support for peer and reciprocal tutoring, and intelligent "matching" of group members. In general, automatic assessment of individual performance in a team environment is limited because of the complex nature of doing plan recognition on groups of human participants. However, in some cases it is feasible to provide one-on-one style tutoring to individuals in a team environment. One example appears in Livak (2004) in the form of a cognitive tutor that supports tactical operations in a 3D first-person "shooter" game.

Natural language dialogue. In face-to-face situations, human tutors use a variety of communicative techniques such as body language, gesture, hesitation, intonation, and, of course, dialogue. Because many intelligent tutoring systems avoid the use of natural language (often called "2nd generation systems"), researchers have suggested that improved natural language dialogue may help "close the gap" between human and computer tutors. This is also motivation for research into *pedagogical agents* that attempt to also leverage non-verbal modes of communication such as facial expressions, body language, and so on.

The pedagogical power of dialogue lies in the increased opportunity for interaction it affords for tutoring systems. Dialogue-capable tutoring systems are now showing learning gains over read-only control groups in support of an *interactivity* hypothesis for learning with tutoring systems (VanLehn et. al., in press). Improved natural language understanding techniques and authoring tools are making it possible to understand student utterances well enough to allow systems to respond in productive and realistic ways. Beginners, who have yet to refine their domain vocabularies but are surprisingly consistent in their language patterns, are ideal targets for modern dialogue-based tutoring technology (Lane & VanLehn, 2005).

ITS as a catalyst in the development of a science of learning. An often overlooked benefit of automated tutors is that their behavior can be "dialed" to test specific hypotheses about learning and tutoring effectiveness. This is difficult to do consistently with human tutors. An area where this strength is quite evident is in the study of feedback. Given a "good" ITS, it is usually a straightforward matter to experimentally adjust the frequency, form, and content of feedback messages and test for difference in learning. For example, McKendree (1990) conducted a study comparing feedback types in a geometry proof tutor. The study showed that goal-directed feedback (i.e., forward-looking hints) led to better performance than backward-looking feedback that flagged errors or explained why steps were incorrect. Studies like this one play an important role in the search for methods of effective instruction and guided learning. At the Pittsburgh Science of Learning Center (PSLC), intelligent tutoring technology is being leveraged (alongside a host of traditional learning science approaches) to address a broad range learning issues and develop a robust theory of learning.[1]

Moving forward: Underrepresented areas of ITS research

To address the future training and educational needs of the Army, it is likely that certain areas of research will need heightened focus from the ITS, AI, and learning science research communities. In this section, we begin with a summary of Loftin et. al.'s (2004) recommendations, then move into brief discussion of areas that appear to be gaps in the ITS community's overall research outlook.

Summary of recommendations of Loftin et. el. (2004)

Loftin et. al. (2004) conducted a large-scale review of ITS research including analysis of the current U. S. Army training requirements, review of non-DoD funded ITS research programs, and interviews with ITS experts and TRADOC personnel. Their recommendations target useful results in eighteen months to three years. The recommendations are broad, including basic and applied research, and a call for large scale transition into some existing Army training program. A few of the specific highlights are:

- creation of an ITS *ontology* to organize ITS concepts and facilitate consensus building
- development of a mapping between classes of ITS architectures and application domains
- increased research into pedagogical agents and virtual humans
- further research and prototype tutors for *team training*
- continued development of ITS development and authoring tools

Two domains are highlighted as potential targets for transition of ITS technology into current Army training: *military history* and *battle analysis*. These are domains of importance in Army training and consist of primarily well-defined components that are within the scope of modern ITS technology. A visible and large-scale integration would be an important proof of concept and example for accomplishing ITS transition efforts in the future. If this goal is adopted, it will be important to apply the lessons learned and basic formula from successful instances of transition (e.g., Koedinger et. al., 1997).

[1] More information about the PSLC can be found at www.learnlab.org. It is one of four science of learning centers funded by NSF that all share the common goal of advancing learning research (www.scienceoflearning.org).

Tutoring and assessment in ill-defined domains

Significant progress has been made in the ITS field for well-defined domains such as algebra, physics, and computer programming. In these domains, the boundaries between right and wrong are crisp – given a model of expertise, it is usually straightforward to immediately assess an action as correct or incorrect. Generating appropriate tutor feedback messages also benefits from this clarity. A good number of domains, including many with particular relevance to Army training, resist such clean models of expertise. These are often described as *ill-defined domains* and have received less attention. If a domain is clearly not well-defined and seems to involve choices that are not obviously right or wrong, there are two possibilities regarding its true nature:

1. The domain *appears* to be ill-defined, but is in reality well-defined – it simply requires further "unpacking" through cognitive task analysis or other forms of analysis.
2. The domain is in fact ill-defined, consisting of instances of subject matter expert disagreement and elements of subjectivity in evaluation criteria.

Successful intelligent tutors have been built for domains like legal reasoning, art interpretation, cultural awareness, and database design. Even with prototype systems like these, the extent to which modern ITS technologies are applicable to ill-defined domains remains an open question. Some technologies are generally robust enough to handle the lack of domain clarity – for example, Bayesian modeling is agnostic to what nodes represent and robust with regard to how updates are made. Others tend to be less of a fit. For example, model tracing algorithms often rely on the ability to evaluate actions as correct or incorrect. In general, as research on ill-defined continues, the fit of existing capabilities to the unique demands of ill-defined domains will become more clear.

Two basic areas of related research appear to need immediate attention. The first is development of detailed accounts of expert behavior in ill-defined domains through cognitive task analyses and other knowledge acquisition tasks. An exemplar of this kind of research Sternberg et. al.'s (2004) influential research on leadership, practical intelligence, and tacit knowledge. The second area is to understand how human experts perform *assessment* in ill-defined domains. We require detailed accounts of the decision-making processes instructors use to understand, classify, and give feedback to students in ill-defined domains. Game environments provide an ideal context in which to collect assessment data and begin to answer these questions. Raybourn et. al. (2005) has adopted this approach by developing a multi-player game for negotiation training that allows human instructors to observe events, log assessments, and provide guidance.

Serious games and narrative learning environments

Sternberg et. al. (2004) found that leadership expertise is bound to experience. In order to accelerate the development of leadership skills, then, it is argued that *experiential* and *narrative-based learning* (i.e., the use of story) should play a role to begin to build a foundation of experiences in learners. This suggests interactive story-telling environments could play an important role in the next generation of leadership training tools. A particularly appropriate context for participating in narrative and practicing skills is provided by modern gaming environments. A relatively new area of research and commercial application, known as *serious games*, attempts to combine realistic simulations of real-world phenomena with the motivational and goal-based features of games.

Frequently, serious games are built with education and training goals in mind from the beginning (e.g., Raybourn, 2004). Unfortunately, there is a conspicuous absence of rigorous evaluations for learning in serious games, so it is not clear yet if expected learning gains are simply not being realized or if more research needs to be done. It is possible that serious games are suffering from the same problems that plagued discovery learning environments (Kirschner, et. al., 2006), and so the role of intelligent tutoring represents an important area of future research to provide the necessary guidance for learners. Several systems represent early attempts to merge these two technologies. Murray (2006) has integrated intelligent tutoring with tactical planning and mission execution, Core et. al. (2006) have built a tutor to support acquisition of interpersonal skills and cultural awareness, and finally, Johnson et. al. (2006) provide a coach for players of a 3d game that teaches conversational Arabic and cultural awareness.

A final area of research that has received very little attention from the ITS community lies in the intelligent manipulation of the simulation itself to achieve pedagogical goals. For example, difficulty changes have been an important component of commercial games for years to enhance entertainment value. An interesting research challenge is presented by exploring the space of difficulty and game behavior adjustments to see how they might be "dialed" to promote learning. Because the best tutors "know when not to" (they intervene only when necessary), this kind of "stealth" tutoring is a particularly appealing path for future research.

Metacognitive tutoring

Domain experts tend to have highly developed metacognitive skills that evolve over time and accumulate with experience. These skills include planning, reflection, and reasoning about hypothetical situations. Although a number of recent tutoring systems have targeted metacognitive skills, more work needs to be done, specifically for ill-defined domains. Metacognitive skills seem to play *more* of a critical role for skills such as critical thinking and decision making. Other areas of AI research play significant roles in tutoring systems that target metacognitive skills, such as natural language processing, dialogue systems, and commonsense reasoning. Continuing fundamental research into ill-defined domains should include detailed analyses of how human tutors operate in them. This will be an important step into understanding the role of reflection and how to scaffold productive introspective skills for improvement and growth.

Automatic detection of unproductive behaviors

Because learners often lack necessary background knowledge and metacognitive skills, it is common to see them display behaviors that are unproductive. It is important to provide guidance at these times so that productive learning can resume as soon as possible. Given the incredibly limited windows of time available for training in many contexts, it is critical to minimize unproductive time. Two categories of unproductive behaviors have been pursued. The first is *gaming behavior*, defined as ways learners will misuse a learning environment to make progress and achieve apparent goals (Baker et. al., 2006). The most common form of gaming tutoring systems is when learners overuse demand help facilities. If students learn that if they ask for help 5 times in a row then get an answer, for example, then they will often rapidly cycle through the less helpful hints to get to the "give-away." A second form of unproductive behavior that has almost no significant research effort is *floundering*. When students are stuck in a learning environment they will often start "trying things" in the hopes they will do something that helps.

This often involves pulling down menus, clicking on buttons, and so on. It is dangerous when they are successful because on the surface, they will succeed. However, since no domain knowledge is involved in using the strategy, it is has no hope of producing desirable learning outcomes.

Conclusions and outlook for intelligent tutoring

Much like the AI community in general, ITS research has resulted in a large body of algorithms and techniques that can be applied to educational problems. The specific advantage of AI-based educational software is its ability to represent domain knowledge and scaffold learning in interactive and deep ways that are not possible in other kinds of learning environments. Learning science research has shown the importance of guidance for effective and efficient learning. For situations when human guidance is unavailable (e.g., while at home) or of limited availability, ITS techniques can help fill this void by giving automated feedback to learners. As the Army revises training practices to reflect science of learning findings (Clark, 2004), it will be essential to include the provision of timely and relevant feedback in computer-based simulations and games for training.

Intelligent tutoring systems have made significant strides in the last few decades in well-defined domains. This paper has suggested an increase in focus on the problems posed by tutoring in ill-defined domains, like leadership and interpersonal skills, will be necessary if the educational and training demands of the Army are to see similar benefits from ITS technology. Specifically, ill-defined domains present research challenges in knowledge representation, learner modeling, capturing expertise, and in authoring. Many believe that serious games provide a motivating and interesting context for learning. The role of tutoring in serious games, both in individual and team contexts, needs to be explored and better understood. Research in dialogue systems have begun to show promise in the context of intelligent tutoring, so this momentum should also continue. If needs such as these are fueled now, there is little reason to believe that early successes of ITS will not be repeated for the new classes of emerging educational and training challenges facing the U. S. Army.

References

Aleven, V., McLaren, B., Sewall, J., Koedinger, K. R. (2006). The Cognitive Tutor Authoring Tools (CTAT): Preliminary evaluation of efficiency gains. In M. Ikeda, K. Ashley, & T. W. Chan (Eds.), *Intelligent Tutoring Systems 2006, LNCS 4053* (pp. 61-70). Berlin: Springer-Verlag.

Anderson, J. A., Corbett, A. T., Koedinger, K., and Pelletier, R. (1995). Cognitive Tutors: Lessons Learned. *The Journal of the Learning Sciences, 4(2),* 167-207.

Baker, R. S. J. d., Corbett, A. T., Koedinger, K. R., Evenson, E., Roll, I., Wagner, A. Z., Naim, M., Raspat, J., Baker, D. J., & Beck, J. (2006). Adapting to when students game an intelligent tutoring system. *Proceedings of the 8th International Conference on Intelligent Tutoring Systems*, 392-401.

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.

Bransford, J. D., Brown, A. L., & Cocking, R. R. (Eds.). (2000) *How people learn: Brain, mind, experience, and school.* Washington, DC: National Academy Press. http://www.nap.edu/openbook/0309065577/html/

Chi, M. T. H., Siler, S. A., Jeong, H., Yamaguchi, T., & Hausmann, R. G. (2001). Learning from human tutoring. *Cognitive Science 25,* 471-533. Retrieved from http://www.pitt.edu/~chi/papers/image3.pdf

Core, M. G., Lane, H. C., van Lent, M., Gomboc, D., Solomon, S., & Rosenberg, M. (2006). Building Explainable Artificial Intelligence Systems, *Proceedings of the 18th Conference on Innovative Applications of Artificial Intelligence*, Boston, MA.

Clark, R. E. (2004). *Design document for a guided experiential learning course*. Final report on contract DAAD 19-99-D-0046-0004 from TRADOC to the Institute for Creative Technologies and the Rossier School of Education http://www.usc.edu/dept/education/cct/publications/clark_gel.pdf

Jermann, P., Soller, A., & Muehlenbrock, M. (2001). From mirroring to guiding: A review of the state of the art technology for supporting collaborative learning. In P. Dillenbourg, A> Eurelings, & K. Hakkarainen (Eds.) *European Perspectives on Computer-Supported Collaborative Learning, Proceedings of the First European Conference on CSCL* (pp. 324-331). Universiteit Maastricht, Maastricht, The Netherlands.

Johnson, W. L., Vilhjalmsson, H., & Marsella, S. (2005). Serious games for language learning: How much game, how much AI? *Proceedings of Artificial Intelligence in Education 2005*, Amsterdam, The Netherlands.

Kirschner, P., Sweller, J., and Clark, R. E. (2006). Why minimally guided learning does not work: An analysis of the failure of discovery learning, problem-based learning, experiential learning and inquiry-based learning. *Educational Psychologist, 41(2),* 75-86.

Koedinger, K. R., Anderson, J. R., Hadley, W. H., & Mark, M. (1997). Intelligent tutoring goes to school in the big city. *International Journal of Artificial Intelligence in Education, 8,* 30-43. Retrieved from http://act-r.psy.cmu.edu/papers/231/jaied97.pdf_2.pdf

Lane, H. C. & VanLehn, K. (2005). Teaching the tacit knowledge of programming to novices with natural language tutoring. *Computer Science Education, 15(3),* 183-201. http://people.ict.usc.edu/~lane/papers/lane05-CSE-NLT.pdf

Livak, T. (2004). *Collaborative Warrior Tutoring*. Unpublished masters thesis, Worcester Polytechnic Institute, Worcester, MA. Retrieved from http://www.wpi.edu/Pubs/ETD/Available/etd-0824104-152548/

Loftin, R. B., Kenney, P., & Mastaglio, T. (2004). Outstanding research issues in intelligent tutoring systems. Prepared under contract N61339-03-C-0156 for the U.S. Army Research, Development and Engineering Command. Retrieved from http://www.mymicsurveys.com/site/files/pub_4.pdf

McKendree, J. (1990). Effective feedback content for tutoring complex skills. *Human Computer Interaction, 5,* 381-413.

Merrill, D. C., Reiser, B. J., Ranney, M., & Trafton, J. G. (1992). Effective tutoring techniques: A comparison of human tutors and intelligent tutoring systems. *The Journal of the Learning Sciences, 2(3)*, 277-305.

Murray, B. (2006). Intelligent tutoring systems for commercial games: The virtual combat training center tutor and simulation. In *Proceedings of the Second Artificial Intelligence and Interactive Digital Entertainment Conference,* 66-71. Menlo Park, CA: AAAI Press.

Murray, T., Blessing, S., & Ainsworth, S. (Eds.) (2003). *Authoring tools for advanced technology learning environments: Toward cost-effective adaptive, interactive and intelligent educational software.* Norwell, MA: Kluwer Academic Publishers.

Niemiec, R. & Walberg, H. J. (1987). Comparative effects of computer-assisted instruction: A synthesis of reviews. *Journal of Educational Computing Research, 3,* 19-37.

Raybourn, E. M., Deagle, E., Mendina, K., & Heneghan, J. (2005). Adaptive thinking & leadership simulation game training for special forces officers. *Proceedings of Interservice/Industry Training, Simulation, and Education Conference.* Orlando, FL.

Sawyer, R. K. (Ed.) (2006) *The Cambridge Handbook of The Learning Sciences.* Cambridge University Press. http://www.cambridge.org/catalogue/catalogue.asp?isbn=0521607779

Shute, V. J. & Psotka, J. (1996). Intelligent tutoring systems: Past, present, and future. In D. H. Jonassen (Ed.), *Handbook of research for educational communications and technology* (pp. 570-600). New York: Simon & Schuster Macmillan.

Sternberg, R. J., Forsythe, G. B., Hedlund, J., Horvath, J., Snook, S., Williams, W. M., Wagner, R. K., & Grigorenko, E. L. (2000). *Practical intelligence in everyday life*. New York, USA: Cambridge University Press.

VanLehn, K., Lynch, C., Schulze, K., Shapiro, J.A., Taylor, L., Treacy, D., et. al. (2005). The Andes physics tutoring system: Five years of evaluations. In G. McCalla & C. K. Looi (Eds.), *Artificial Intelligence in Education*. IOS Press.

VanLehn, K., Graesser, A. C., Jackson, G. T., Jordan, P., Olney, A., & Rose, C. P. (in press). When are tutorial dialogues more effective than reading? *Cognitive Science.*

Youngblut, C. (1994). Government sponsored research and development efforts in the area of intelligent tutoring systems. Institute for Defense Analyses (IDA Paper P-3003). Retrieved from https://www.spider.hpc.navy.mil/html-docs/its-ida/