Intelligent Tutoring Systems: A Literature Synthesis

Nathaniel Raley

1 Review Objectives

My experience working as a one-to-one tutor in different subjects and with students of various ages has given unique breadth to my understanding of *tutoring* as an indispensible tool for education. This subjective claim enjoys firm support in the literature; in an often-quoted study, respected educational psychologist Benjamin Bloom reports that the average student tutored one-to-one performed two standard deviations better than students who were taught in a conventional classroom setting, a finding which held across grade-levels and subject matter (Bloom, 1984). But the benefits of this instruction would seem to be the province of the wealthy few, those who could afford to hire full-time tutors for their children.

Thankfully, technology is moving in a direction that promises to change all of that. Having revolutionized science, commerce, and communication in a matter of decades, computers are poised to do the same in education. And with almost 2.5 billion people across the globe connected by the internet (internetworldstats, 2012), the role of the internet in transforming education may well rival that of the printing press. It has already changed how educational material is produced by reducing its cost and improving its efficiency, to say nothing of its more immediate effects on students. Now the world's largest repository of knowledge is for most only a mouse-click away, moving education down from ivory towers and unfettering it from constraints of space and time. It has certainly crept its way into the classroom: in 2009, 97 percent of teachers had one or more computers located in the classroom every day (National Center for Education

Statistics), and the ratio of students to computers in the classroom was 5.3 to 1 (and this was four years ago)! The flip-side of this technological surge is that we have created a "knowledge economy" that requires its workers to keep pace with rapidly changing demands; accordingly, job skills in many fields (particularly information/technology) will need to be updated continuously. These sorts of issues, the dream of "a teacher for every student," and support for efficient lifelong learning may be effectively resolved through the development and implementation of computer-based Intelligent Tutoring Systems (ITSs).

I am pursuing this topic for several reasons:

- I am interested in the nexus of psychology, education, and artificial intelligence; all three fields must collaborate to create effective ITSs, and I want to understand the extent to which this collaboration has occurred (and whether or not extant ITSs are pedagogically sound with respect to the Educational Psychology literature)

- I want to understand the architecture of these systems and the theories from artificial intelligence, statistics, and cognitive psychology on which they are based; I have found an open source (linux-based) cognitive tutor called GnuTutor (based on AutoTutor) that I plan on exploring

- I believe that this technology has a lot of educational potential, and I feel that this really represents an increasingly "Current Topic" in Learning and Instruction; I want to document some of the more successful of these systems, their successes and their shortcomings, to get a clearer picture of *the state of the art*, with discussion of topics like machine learning, collaborative environments, and web-based tutors

Intelligent Tutoring Systems are the progeny of artificial intelligence and education; the goals of ITS development are as lofty as they are numerous. Some early statements of these goals were "to replicate the one-to-one interaction between a tutor and a learner" (Dede, 1986), to reason about a student's knowledge, monitor the student's progress, and adapt the teaching strategy to the student's individual learning pattern (Woolf, 1987), to assign problems to students on an individual basis, monitoring students' solution steps, providing context-sensitive feedback, and implementing a mastery learning criterion (Anderson et al., 1995). More recent goals of ITS have included scaffolding student understanding of complex content, facilitating self-explanation for deeper understanding (Aleven and Koedinger, 2002), creating engaging and authentic learning contexts (Kindney and Pucket, 2003; Dede 2009), reasoning about a student's affective state (mood, motivation, interest) and discovering links between behavior (time spent, number of hints) and hidden variables (attitudes, goals, self-efficacy) (Woolf, 2008), representing information in various ways to encourage generalization and transfer of knowledge (Koedinger et al, 1997; Anderson et al., 1995)... the possibilities are conceptually limitless. Intelligent Tutoring Systems are what McArthur and Lewis (1998) have called the "holy grail of teaching technology."

In the late 1960s and into the 1970s, Computer-Assisted Instruction (CAI) emerged in the wake of advances in computer science and the burgeoning field of artificial intelligence; CAI projects were initially limited to programmed instruction, functioning merely as an educational tool for content delivery. With advent of the personal computer in the late 1970s and early 1980s, it became clear that computers could themselves act as teachers. Research in Cognitive Psychology and Artificial intelligence laid the foundation for these new Intelligent Tutoring Systems, an early example of which was LISPITS in 1983 which enjoyed some breakthrough successes teaching students the LISP programming language (Corbett & Anderson, 1992). Now, successful implementation of ITSs is fairly well-documented; some examples are AnimalWatch (designed to motivate 6^{th} graders to use math in a problem-solving context, embedded in an engaging narrative about endangered species (Arroyo et al. 2001, 2004; Beal 2001) and Pump Algebra Tutor (PAT, a full-year algebra course for 12–15 year-olds developed at Carnegie Melon University which dramatically improved albegra skills and problem solving compared to traditional albegra classrooms (Koedinger et al., 1997). There is also an increasing movement toward the development of ITSs based on conversational dialogue (AutoTutor, Atlas, Why2, Andes).

According to Woolf (2008), an Intelligent Tutoring System differs from other computer-based education in at least seven distinct ways:

- **Generativity** – the ability of the system to generate problems, hints, and help customized to students learning needs (e.g., basing problem presentation on student learning, infering a student's solution plans from a partial sequence of actions)
- **Student Modeling** – the ability of the system to represent and reason about a student's current knowledge and learning needs and to respond by providing instruction (e.g., track student acquisition/retention to generate timely individualized problems and hints)
- **Expert Modeling** – the ability of the system to represent and reason about expert performance in a domain; representing the topics, concepts, and processes of a given domain (e.g., math knowledge represented as if-then production rules, then student solutions generated as steps and missteps; then identify students' strengths and weaknesses relative to the rules they used)
- **Mixed Initiative** – the ability of the system to initiate interactions with a student as well as to respond usefully to student-initiated interactions; ability for either the student or the tutor to take control
- **Interactive Learning** – the ability of the system to provide learning activities that require student engagement and are authentically contextualized and domain-appropriate (e.g., immersive interfaces)
- **Instructional Modeling** – the ability of the system to change teaching mode based on inferences about a student's learning; how a tutor modifies its guidance

for each student (requires input from the student)

- **Self Improving** – the ability of the system to monitor, evaluate, and improve upon its own teaching performance based on the system's experience with previous students; often implemented through machine-learning and datamining techniques that evaluate prior students' learning experiences, judge which interventions were effective, and use this information to inform its responses.

The topic of educational technology typically brings to mind rueful images of souped-up whiteboards or powerpoint presentations replete with flashy graphics and over-the-top multimedia; given many missteps and wrongheaded applications of technology to education, it is easy to see how it got this reputation. However, I don't see that happening for ITSs; I believe that, when used in traditional classrooms with caring teachers, this can provide all the advantages of individualized instruction at a fraction of the cost. I see them speeding through topics that a student already grasps, focusing on topics that are troublesome, and never losing patience. I believe that, among other aims, this technology has the potential to accomodate individual differences and respond effectively, to track student collaboration and identify student contribution in groups, and to motivate bored, tuned-out students; in short, I think that its addition to instructional design would go a long way toward optimization of the learning process.

2 Introduction and Background

The central premise of educational technology is that instruction can be improved through thoughtful application of experimental findings from educational psychology and cognitive science to educational software informed by advances in artificial intelligence; because there are individual differences from student to student, each will learn more effectively and efficiently when the material is customized and individualized to suit his or her specific situation. Woolf (2009), a leading researcher and developer of

Intelligent Tutoring Systems, states that the goal of incorporating artificial intelligence in education is "to match the needs of individual students by providing [individually optimal] representations of content, [individually optimal] paths through material, and [individually optimal] means of interaction.." in ways that are "...pedagogically sound." What *exactly* all of this means will be explained hereafter, as will an objective evaluation of this technology, its successes, and its shortcomings.

The potential value of computers in education was recognized before computers themselves became a reality, but the advent of the personal computer in the mid-eighties made this potential evident, the spread of the world wide web made it clear, and today's culture of mobile computing (laptop computers, smart phones, tablets, e-readers...) have made it downright obvious. All of this innovation has wrought sweeping changes to modern societies and economies, increasing the demand for knowledge-workers whose job skills will need to keep pace with these rapid changes as they continue; the worker of the future will not only be required to have a broad knowledge base, but also to be learning constantly and with little outside support. It can be argued that the confluence of the internet, artificial intelligence, and cognitive psychology creates an unprecedented opportunity to improve education for people everywhere. Computers have revolutionized science, communication, and commerce in only decades; are they poised to do the same in education?

"Supplying students with their own automated tutor, capable of finely tailoring learning experiences to students' needs, has long been the holy grail of teaching technology"
–McArthur and Lewis, 1998

Intelligent Tutoring Systems (hereafter, ITS) are a subset of "Intelligent Educational Software," which is defined more generally as software that uses intelligent techniques to model, reason about, and respond to learners (Woolf, 2009). This software is itself a subset of the techniques encompassed by the inclusive term "Computer-Based Instruction". ITSs defy succinct definition; one goal of the present paper is to reconnoiter the vast research landscape well enough to map these systems out with reasonable accuracy.

Master teachers know the domain to be taught and they know how to teach it; they have amassed formidable experience over many years of teaching, working with multiple students simultaneously and learning about their knowledge. They use various instructional strategies to work dynamically with students of differing abilities, experiences, cultural backgrounds, likes/dislikes, personalities, goals, moods, etc. They give hints, explanations, and examples; they encourage reflection, self-regulation, and motivation; they consider timing, sequencing, and learning goals; they are cognizant of troublesome knowledge, typical errors, and incorrect strategies; and they are mindful of individual student characteristics like those listed above (Shute, 2006). They may intervene to modulate a student's self-confidence, to elicit curiousity, to challenge their students, or to allow students to feel in control. Indeed, during one-to-one tutoring, human teachers devote at lease as much time to reasoning about their student's emotion as to their cognitive goals (Lepper and Gurtner, 1989).

Obviously, ITSs require tons of encoded knowledge if they are to be successful; this includes knowledge about the target domain, the student, teaching practices, and how to "communicate." Before I begin, though, I want to adumbrate the conditions necessary to define an ITS, as well as some of their more important ancillary features. The minimum requirements are that a system is able to represent "domain knowledge" (e.g., ideal solutions of all problems in a given domain) and "student knowledge" (e.g., the extent to which a student is able to solve problems in a given domain, the performance of past students, possible misconceptions, time spent on problems, hints requented, number of errors); the system can then provide instruction or other educational feedback based on these representations, and modify them according to student input (Anderson et al., 1995). Both of these models, and the interactions between them, will be discussed in more detail in the following chapters.

It is important to notice that ITSs are, by nature, student-centered; they model a student's knowledge about a topic and adapt the model over time as understanding becomes more sophisticated (Corbett and Anderson, 2008). They can depict key ideas (including threshold concepts) as well as common misconceptions (troublesome knowledge, refutational texts) in a pedagogically-effective and individually-appropriate sequence. Increasingly, they are able to model student variables beyond their knowledge of a topic, including a student's engagement and affect. What's more, they can now respond in ways that go beyond fixed instruction, such as changing teaching strategy when appropriate, specifically training a student's metacognitive skills, and providing students with emotional and motivational support. The potential of these systems is so great, and their success so unequivocal, that they cannot but alter existing teaching and learning practices. But this is not at all to downplay or otherwise to threaten the teacher's lynchpin role in the classroom; on the contrary, ITSs allow teachers to provide individualized attention to students who need it most and producing objective data for them about student understanding and misunderstanding, engagement and disengagement, satisfaction or dissatisfaction.

2.1 A Brief History of Computers in Education

The history of computers in education is storied and interesting, roughly tracking the vicissitudes of artificial intelligence research. After the 1944 G.I. Bill, which provided free college education to veterans of World War II, went a long way toward democratizing higher education in the United States; it also caused great surges in enrollment, and by the mid fifties university administrators began to seriously worry about logistics of educating the scads of new students. It was clear that computerized automation was effective in increasing factory production... could it not do the same for education? Spurred by the Soviet Union's launch of Sputnik I, the government increased spending more on science and engineering education. The first computer used in education dates back to around this time. In 1959, PLATO was created at the University of Illinois after several unproductive meetings between university administrators, engineers, mathematicians, and psychologists. It was somewhat effective, undergoing many revisions until become defunct in 2006, but the technology available at the time was too rudimentary to realize many of the goals (Shute and Psotka, 1994).

Another big step forward came in the 1970s with the LOGO programming language, developed specifically to encourage students to think rigorously about mathematics by creating a “math land” where students could “play” with mathematical concepts (Papert, 1993). Seymour Papert, the inventor of LOGO, was influenced heavily by Piaget and constructivist learning theory. Indeed, LOGO was designed not only to teach programming and computation concepts but to increase a child’s well being in a culture increasingly dominated by technology: “more important than having an early start on intellectual building, is being saved from a long period of dependency during which one learns to think of learning as something that has to be dished out by a more powerful other...Such children would not define themselves or allow society to define them as intellectually helpless.”

Rapid developments in technology led to equally rapid increases in what came to be called Computer-Based Instruction (CBI). Especially when compared to today’s ITSs, these early systems were very basic and limited in nature: they were frame-based and fixed, with the sequence of problems predefined and inflexible. Students could not ask questions or request hints unless they were specifically pre-programmed by the author. There was early support for viewing simulations, changing parameters, adjusting difficulty levels, and so on, but they were not at all adaptive... they provided the same fixed instructional environment to each student. Even still, CBI was found to be effective! A meta-analysis of several hundred well-controlled studies showed that CBI increased class performance by about one-half standard deviation and decreased learning time by a third; it also improves students’ attitudes and interests through more interactive, enjoyable, and customizable learning (Kulik and Kulik, 1991).

2.2 Intelligent Tutoring Systems

Despite these early successes, “students do not learn by simply pressing buttons, even if the new pages contain animations, images, sounds, video, or other ‘multimedia’. Instead, exercises should preferably involve students in the material and be adaptable to different learning needs” (Woolf, 2009). The first genuine ITS was born out of a Ph.D. thesis project by Jaime Carbonell in 1970; the system he developed, called “Scholar”, enabled students to explore the geographical features of South America and, significantly, provided individualized responses to student statements. However, in its extremely basic incarnation it lacked an expert system; this came later, in 1979, with a system that represented expert approaches to medical diagnoses and guided students through the process (it was aptly named GUIDON).

A watershed moment for ITSs occurred in 1983 with the creation of LISPITS by John Anderson and his colleagues at Carnegie Mellon University (Corbett & Anderson, 1992). LISPITS was built to teach students how to program in LISP, an important language with many applications in the field of artificial intelligence; the tutor required students to enter code one step at a time, but could identify where mistakes were made and could provide specific feedback to students as they worked towards solving the problem. It is first example of what has come to be termed a “Cognitive Tutor,” and was built upon the assumptions of John Anderson’s ACT-R model of human cognition, which greatly influenced the future development of the field (Anderson, 1983). More generally though, Cognitive Tutors are a subset of ITSs predicated on information processing theory; this perspective holds that learning is essentially a computational process, and it relies on the crucial assumption that the humans and computers process information in comparable, if not identical, ways (i.e., they encode, store, and retrieve information). Often, they represent expert solutions for a given task as a series of steps while simultaneously recording a student's input at each step and comparing it to the expert solution. When the student's solution steps deviate from the expert solution steps, the tutor can offer feedback and remedial instruction contingent on the specific step that is at issue. Cognitive tutors were also the first commercially successful ITSs: Carnegie Learning, a company founded by researchers at Carnegie Mellon University, produced the commercial ITS for use in high school mathematics classes. More than 475,000 students in more than 1300 school districts across the United States used this curriculum, or about 10% of the U.S. high school math classes in 2007 (Woolf, 2009).

2.3 Modern ITSs

All too often, the adoption and implementation of well-supported learning theory in actual instruction lags far behind their acceptance in the educational psychology and cognitive science research community; in ITSs this tendency may be even more pronounced (see Section 3). Still though, in this brief history we see researchers incorporating constructivist and information-processing approaches in their development of these systems relatively early on. Still more reassuring, Modern research in ITSs is exploring the implementation of collaborative environments, inquiry learning, and web-based tutors (see Section 2). Beyond learning theory itself, there has been much recent interest in developing natural language tutors that can simulate conversation with the student and interpret student dialogue (Graessner et al. 2001), as well as tutors that can interpret and monitor student affect and motivation, enabling the system to respond in ways that maintain an optimal state for learning (Shute, 2006). The following sections will address the state of the art of ITSs and the degree to which they have support in the learning literature; we will end with a thorough evaluation of their effectiveness.

3 Intelligent Tutoring Systems and Learning Theory

On their own, models of student and expert knowledge can achieve very little; ITSs rely on instructional knowledge to govern the way they respond to students. This knowledge of how to teach is fundamental, determining when and how to intervene based on estimates of a student's knowledge, emotions, and motivation. Many effective teaching strategies are difficult to implement in classrooms, costly and resource intensive, or both. For instance, apprenticeship training requires a 1:3 teacher-to-student ratio, and it has been shown that teachers often fail to modify their instruction to account for new, effective practices; teachers are likely to default to the manner in which they themselves were taught, rather than the way they were trained to teach (Feldon, 2007). Another upsetting fact observed in classrooms is something called the Pygmalion effect, where a teacher's preconceptions of student traits and abilities are better predictors of subsequent evaluations than actual classroom performance (Feldon, 2007). Similar biases, including those related to physical attractiveness, have the same effects. The potential for ITSs to consistently implement these new strategies in an objective and unbiased way is clear, because they can be programmed to do so, but it doesn't stop with consistency. These systems can (or will soon be able to) flexibly alter their teaching strategy, select interventions, customize responses, and motivate students.

However, developing such a system—one that can effectively deliver and flexibly modify instruction—raises many questions (du Bouley and Luckin, 2001). Should programs be modeled on human teaching approaches? For which domains is each approach best suited? What aspect of a given instructional strategy is critical for its success? Also, the idea that no single approach to learning is appropriate for *all situations* or *all students* is a powerful driving force behind ITS development; instructional strategies should be flexibly selected based on considerations of the domain, the nature of the concepts to be learned, and learner variables such as current ability and engagement. Domains with low cognitive processing and highly prescriptive solutions, like algebra procedures, might benefit more from traditional instruction and practice with problem solving, while domains that demand higher levels of processing, such as problems in physics, history, and the social sciences, are often better served by a constructivist approach. Some domains require the student to have more control over the environment, and some that are highly integrative, such as software development or medical training, might be better taught using situated approach.

The sequences and stages of learning, too, are important to consider. For example, initial knowledge acquisition is almost always declarative in nature; if it is knowledge of a skill, then practice with performance will allow for procedural representations of that knowledge. If it is knowledge required to understand a larger concept, then constructivist instruction may be more appropriate, aiding integration and interrelation of disparate concepts. As an added complication, there is often considerable overlap between approaches; to take just one example, Brown, Collins, and Duguid (1989) say that apprenticeship learning "enables students to acquire, develop, and use cognitive tools in authentic domain activity. Learning, both inside and outside of school, advances through collaborative social interaction and the social construction of knowledge." This single-sentence definition requires aspects from many disparate learning theories, including active learning, socially based learning, constructivism, and situated learning. Given these difficulties of categorization, the following groupings are more convenient than precise; they are (1) active learning, (2) collaborative and socially based learning, (3) constructivist and inquiry learning, (4) situated learning, and (5) self-regulated learning. The onus is on me here to provide compelling evidence that learning theory is being prudently and effectively employed in Intelligent Tutoring Systems, and that these systems have the potential to integrate their sundry features into an process that can optimize the learning process.

3.1 Active Learning and Problem Solving

Teaching through problem-solving is an area where ITSs really shine; it's no wonder that some of the most successful instantiations of ITS have been geared toward quantitative domains, which are often taught through complex, multistep problems. Often, problems like these occasion heavy cognitive load for students, especially if the intermediate skills involved have not become procedural and effortless through practice. By breaking solutions down into their constituent steps, making sure the prerequisite concepts have been mastered, and allowing for multiple representations (spreadsheets, points on a graph, equations), tutors can reduce cognitive load. Interestingly, cognitive load can be estimated using only a learners pupil size, allowing a tutor with the proper hardware to respond appropriately (Kaklauskas et al., 2013). A few examples of successful active learning tutors include Andes (Gertner and VanLehn, 2000), Pump Algebra Tutor (Anderson et al., 1995), Wayang Outpost, and Animal Watch (Arroyo et al., 2004).

Worked examples (i.e., providing students with pre-worked problems that emphasize each step in the solution) have been shown to reduce cognitive load for students, particularly low-ability students, in the context of a problem solving task (Sweller, 2006), and ITSs are able to show them step-by-step, highlighting important details (Schwonke et al., 2009) Additionally, they can smoothly implement scaffolded problem solving through the "guidance fading effect," where help is provided at first, but then gradually reduced until the student can perform expertly on their own; in a typical classroom setting, this relies heavily on teachers picking the right point at which to alter their instructional techniques. The above instance of differential outcomes for low-ability vs. high-ability students provided with the same instruction is just one manifestation of the something called the "expertise-reversal effect." This effect is seen in the sequencing of problems and examples as well; Sweller (2006) discussed how examples before problems produced better learning for lower-knowledge learners, while examples *after* problems resulted in superior performance for higher-knowledge learners. We will meet with this effect again soon, in our discussion of feedback.

In an instructional context, feedback has been defined as “information communicated to a learner that is intended to modify the learner’s thinking or behavior for the purpose of improving learning” (Shute, 2006). Feedback during problem solving is crucial; it provides useful information for correcting misconceptions, procedural errors, or inappropriate task strategies. For example, recall the LISPITS programming tutor mentioned in the historical overview above. Corbett and Anderson (1991) reported that learning took up to three times longer when feedback was delayed than when it was given immediately. However, this result could’ve been due to the expertise-reversal effect and the fact that all students using the LISP tutor were complete novices in this domain; while immediate feedback for students with low achievement levels is superior to delayed feedback, delayed feedback has been shown to produce better learning for students who are already achieving at a high level (Anderson et al., 1989). ITSs are uniquely suited to capitalize on phenomena like the expertise-reversal effect; they can facilitate better learning through differential instruction of learners based on demonstrable differences in performance.

Not all feedback is equivalent, and feedback of an effective quality and appropriate quantity can have many additional benefits; effective feedback, for instance, provides details about how to improve the answer rather than just indicating whether or not the response was correct (Bangert-Drowns et al. 1991), and changes the learner’s locus of attention, directing attention to the salient aspects of the task (Kluger and DeNisi, 1996). Interestingly, computerized feedback interventions have been shown to yield stronger effects than non-computerized intervention (Baumeister et al., 1990). Feedback can also have important implications for outcome variables other than direct measures of learning. It can modify a learner’s goal orientation, topic self-concept, and view of intelligence by reinforcing the idea that ability and skill are developed through practice, that effort is crucial, and that mistakes are an important part of the whole process

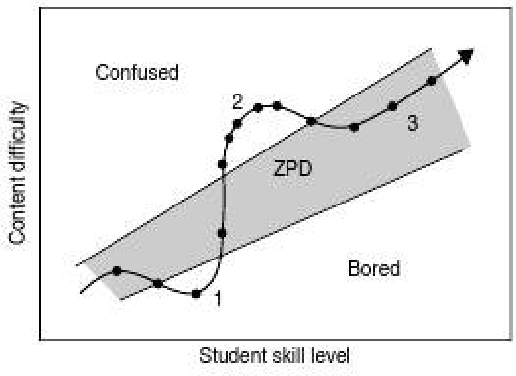
(Narciss, 2004).

3.2 Collaborative and Social Learning

A central tenet of Vygotsky's theoretical work is the idea that every aspect of an individual has its roots *outside of* the individual. "Every function in the child's cultural development appears twice: first, on the social level, and later, on the individual level; first between people (inter-psychological) and then inside the child (intra-psychological). This applies equally to voluntary attention, to logical memory and to the formation of concepts. All the higher functions originate as actual relationships between individuals" (Vygotsky, 1978), suggesting that cognitive skills and patterns of thinking are primarily the product of the sociocultural milieu in which an individual lives, and that the individual's personal history as well as the history of the society and culture, heavily influence learning and thought. Given the of emphasis given to the individual here, many of Vygotsky's more collectivist writings have fallen on relatively deaf ears in the West.

However, one idea of Vygotsky's has enjoyed wide acceptance in the West is the "Zone of Proximal Development" (ZPD), the notion that at any given time, each student has a potential capability that can only be realized through external "scaffolding" in the form of guidance or collaboration by a more-knowledgeable other; when learning is viewed as successive transitions between knowledge states, the purpose of teaching is accordingly to facilitate the student's traversal of the space of knowledge states (Wenger, 1987). Clearly, the optimal conditions differ for each learner and each context, and again, wherever troublesome interactions like these rear their head, ITSs are able to flexible accommodate them. Some tutors have been built with the express purpose of keeping students in the ZPD, maintaining a level of content difficulty that is *just beyond* their current skill level (Figure 1). Under these conditions and given timely guidance, these ITSs prevent confusion and impasses on the one hand, and boredom on the other. An example of such a tutor is EcoLab, a social interaction tutor that dynamically adapted its help and activities to a learner's collaborative capability, thereby ensuring that each student was extended beyond what he or she could achieve individually (du Boulay and Luckin, 2001).

Figure 1: *Schematic of how an ITS might maintain the ZPD, adpated from Woolf (2009)*



Another extremely successful ITS based on social learning is the US military's program DARWARS. This inexpensive, web-based simulation system is based on collaborative multiplayer games that take place in virtual worlds (Chapman et al., 2004). It provides immersive practice for the coordinate of team skills and gives feedback at both the individual and group level. This program has enjoyed wide application by our armed forces; in 2006, more than 20,000 individuals were trained on it. The potential for collaborative learning in online environments such as Massively Open Online Courses (MOOCs) is promising, but this territory remains relatively uncharted.

3.3 Constructivist Approaches and Inquiry Learning

A quote from Perkins (1991) will provide a suitable introduction to this section on constructivist approaches in ITSs: "Information processing models have spawned the computer model of the mind as an information processor. Constructivism has added that this information processor must be seen as not just shuffling data, but wielding it flexibly during learning—making hypotheses, testing tentative interpretations, and so on." Because learning outcomes are not always predictable, instruction should foster rather than control the learning process. Certain features of intelligent tutors enable the purposeful construction of knowledge; however, few intelligent tutors have fully implemented this perspective. One issue is that because constructivism promotes open-ended learning experiences, learning cannot be measured easily and may not be the same for each learner. Below are a few examples of constructivist features that seemed to warrant placement this category rather than another.

A notable success story for constructivism in ITSs is the Rashi tutor; this medical-training tutor invites students to diagnose illnesses by interviewing patients about their symptoms. It imposed no constraints concerning the order of activities; students were free to explore images, ask questions, and collect evidence in support of their hypotheses. Students could then test their hypotheses to discover principles on their own, and they could reach unique conclusions. It is clear form this example that constructivist ITSs share many principles with "situated" ITSs, such as realistic settings, authentic problems and contexts, and evaluation integrated with the task. Accordingly, these ideas will be dealt with in that section at greater length. In other inquiry-based tutors, the expertise-reversal effect has cropped up, calling for more flexibility of instructional strategy in these types of tutors. Several inquiry-based tutors were found to be more ffective for high-apptitude students and less effective for low-aptitude students (Woolf, 2009).

Self-explanation can be viewed through the lens of constructivism or that of self-regulation. Because it taps the generation effect, where students benefit from generating their own personally memorable explanations, I have chosen to include it in this section on constructivist approaches (Hausmann & VanLehn, 2010). Self-explanations occur when students attempt to explain, either implicitly or explicitly, their solution process for a particular problem, and it has successfully been implemented in ITSs. Aleven and Koedinger (2002) show that students who explained their steps during problem-solving practice with a geometry ITS learned and understood more than their control group, who experienced the same ITS except without the self-explanation features.

Self-explanations have certain features in common with "Socratic learning," which has served as the basis for several effective ITSs. Socratic learning centers on the idea that teachers do not need to put ideas *into* students, but rather draw the ideas *out of* students, encouraging them to articulate explanations through dialogue. An early example of an ITS based on a Socratic learning techniques was the WHY tutor (Stevens and Collins, 1977), but their rigid question-answer format has made them less attractive to modern developers. A key feature of Socratic tutors is their ability to cause students to confront their own understanding, or lack thereof. In this way, they serve as something of an interactive 'refutational text', text that directly challenges a student's preconceptions, thereby prompting a restructuring of alternative conceptions and occasioning deep conceptual change (Mason et al., 2008).

Apprenticeship learning typically features an expert who monitors performance, provides advice, and supports multiple legitimate paths to solutions; these experts typically do not engage in explicit tutoring, though they may "scaffold" instruction, providing initial support for the problem-solving process and then handing responsibility over to the student (Brown et al., 1989). Apprenticeship training is familiarly seen with pilots, athletes, physicians, and skilled tradesmen. The idea is to enable students to develop robust mental models through experience with realistic learning conditions; working alongside an expert, these students come to reproduce the requisite actions on their own. An early example of such a tutor was the Sophisticated Instrument Environment (SOPHIE), which assisted learners in learning to troubleshoot faulty electronic equipment (Brown et al., 1982). Apprenticeship learning will be touched on again later from a situated learning perspective.

3.4 Situated Learing and Immersive Interfaces

Situated learning theory states that learning is a function of the activity, context, and culture in which it occurs; all that people do is situated in their role as a member of a community (Lave and Wenger, 1991). Accordingly, social interaction within an authentic context is necessary for learning because students are embedded in a "community of practice" that demonstrates the beliefs and behaviors to be acquired; this is to be contrasted with learning that is abstract, or takes place out of the context in which it is to be used. To effectively provide for situated instruction, assessment must be an integrated part of the learning environment; it must take place within, and not after, the task.

Many situated tutors have been built for military training and other high-risk occupations.The US military is one of the largest investors in electronic training like ITSs (Woolf, 2009); these forms of training are cheaper, faster, more consistently available to enable continuous practice, and more accessible to avoid impracticaities ofl travel. By training in immersive environments that replicate and contexualize their duties, soldiers can acquire skills that previously came only with actual tactical experience. For example, the 'Operator Machine Interface Assistant, provided training for helicopter pilots in the US Navy by simulating the operation of a control console, a mission display, and the flight environment itself (Stottler, 2003). Some, like Steve (the Soar Training Expert for Virtual Environments) go further and include animated pedagogical agents that interact with trainees in an immersive virtual reality environment (Johnson et al, 2000). These agents were 3D and spoke to trainees as they moved around to view the demonstration from different perspectives The agent taught them how to perform tasks in this authentic environment (high pressure air compressor room aboard a Navy surface ship), demonstrated physical tasks, and integrated demonstrations with commentary. NASA, too, has used virtual reality simulations in their ITSs (Woolf, 2009).

Another excellent application of situated learning techniques in ITSs can be found in the domain of language training. Adult learners often have difficulty acquiring even a basic working knowledge of a foreign language, usually because language immersion is impractical if not impossible for most. The Tactical Language Tutor taught thousands of US military personnel to communicate in Iraqi arabic in an effective and culturally-sensitive way by simulating immersion.

Some ITSs feature apprenticeship environments; their efficacy relies less on the guidance of a skilled expert and more on the affordances of a simulated envionment. For example, a program called Sherlock simulated the structure and function of an circuit board removed from an F-15 airplane and allowed trainees to select and use the appropriate measurement tools in a simulated environment (Lajoie & Lesgold, 1992). Sherlock provided authentic problems based on empirical studies and it only provided help to avoid a total impasse. After completing a complex job-related task, which could take upwards of an hour to figure out, the trainees reviewed their actions step by step, and could receive information about the best solution at each step, and why each step was taken.

3.5 Self-regulated Learning and Metacognition

Supporting metacognition during instruction can improve robust learning of domain knowledge (Koedinger et al., 2009). Several corollary goals of training metacognition include improving student's metacognitive behaviors and domain-level learning both during *and after* the intervention. Several ITSs have been built in order to explore the extent to which metacognitive skills can be trained; these include adaptive features such as support for self-explanations (Aleven & Koedinger, 2002), support for error self-correction (Mathan & Koedinger, 2005), methods of reducing a student's "gaming" or exploitation of the system (Baker et al, 2006), and teaching appropriate help-seeking skills (Roll et al., 2007) all in authentic classroom environments. The tutors that implemented metacognitive support for error correction, self-explanations, and reductions in "gaming the system" all led to improvements in domain learning compared to otherwise identical ITSs lacking these components, while four of the tutors improved transfer of the metacognitive skill in question to other tasks, though this effect was much diminished in delayed post tests. The authors note that student affect seemed to play quite a large role in the relative efficacy of these interventions.

3.6 Affect and Motivation

Though the idea is in its relative infancy, some ITSs have already enjoyed success in modeling student affect and motivation. If tutors could reliably detect and smoothly respond to student emotion, they have the power to leverage the central role of affect in producing learning outcomes. It is well known that emotion is completely bound up with cognition in guiding human behavior such as memory and decision making (Woolf, 2009). It shows up in educational settings all of the time: anxiety and depression can paralyze a student's ability to hold task-relevant information in working memory, and as such students experiencing these emotions tend not to properly assimilate information (Goleman, 1996). On the other hand, mildly positive affect has been shown to improve negotiation, generosity, social responsibility, and motivation to learn (Burleson, 2006). In addition to student affect, motivation and self-confidence have been shown to mediate learning. A student's response to difficult tasks depends in no small part on their goal orientation; students with mastery orientations toward academic achievement—those whose desire is to learn for the sake of learning—persevere more in these situations (Dweck, 1986).

Affective and motivational variables can be detected using computer hardware such as metabolic sensors (e.g., cameras, posture sensors, skin-conductance, etc.) or inferred through computer software that can examine logged data of student behavior (e.g., time spent on a question, number of problems seen or attempted, number of hints selected). Indeed, motivation can be quantified in this way with relative ease, even in the absence of sophisticated hardware; by watching videos of student-tutor interactions, de Vicente and Pain (2000) linked motivation to directly observable variables and developed 85 rules that ITSs can use for inferring student interest, satisfaction, control, confidence, and effort. To be sure, developments in hardware can improve upon a software tutor's ability to detect the niceties of affect. Information about student attentiveness, engagement, boredom, and frustration can be deduced from the state of a learner's gaze using eye-tracking strategies (D'Mello et al., 2012).

One example of an ITS that has successfully modeled student affect is Wayang Outpost, a web-based system that helped prepare students for high-stakes testing (Arroyo et al., 2004). This program measured student motivational variables such as whether students thought that they had learned and whether or not they liked the tutor with 80-90% probability, and this was achieved in the absence of auxiliary hardware (Arroyo et al., 2005). A still more striking example of this was demonstrated by D'Mello et al. (2008) using AutoTutor. This tutor was able to significantly predict the affective states of bordeom, flow, confusion, and frustration based on conversational features extracted from student interactions with it; in the same study, it was shown that the tutor's detection of affect was on par with that of novice judges but lower than that of trained judges.

3.7 How are ITSs contributing to the literature on learning/instruction?

In addition to their obvious didactic applications, this technology has utility as a research tool as well; ITSs provide a unique proving ground for the validity of theoretical claims in educational psychology. They are flexible and predictable, making them useful for revealing the truth of cognitive function (Anderson et al., 1984). Moreover, tons of educational data is being collected on learners through learning management software, distance learning courses, MOOCs, and other educational web applications. All of this data is ripe for analytics and is a booming research interest; what little has already been done in educational data-mining has contributed insights useful in the field of educational psychology. Khan Academy, an online learning environment encompassing many subjects, tested data collected from student logs to see whether including different kinds of motivational messages above the problems the students were solving had any effect on performance. They found that generic, positive messages ("this might be a tought problem, but we know you can do it!") had no effect, while messages that reinforced an incremental view of ability ("remember, the more you practice, the smarter you become!") significantly improved performance. Similarly, Coursera, a website providing access to free, online university courses, stumbled upon an insight. They had implemented a feature meant to encourage students by sending them reminders of due dates and upcoming assignments, but it ended up having the opposite effect and led to a drop in retention, with students reporting that they felt "harassed." Coursera is also able to find clusters of distinct types of learners: they recently discovered that there is also a group of students who complete all of the homework assignments without watching any of the lectures (The beauty of data-mining is that sample size is rarely an issue, and when a pattern or trend is discovered, you can test it with relative surety (BBC Future, 2013).

4 A Weighing of the Pros and Cons

Bloom's (1984) frequently cited study was the first to demonstrate the unequivocal benefits of one-to-one tutoring over alternative instructional methods. Therein, he reported that student achievement after traditional classroom instruction was about two standard deviations lower than student achievement based on individual tutoring. Intelligent tutors have likewise been shown produce improvements over typical classroom instruction and can effectively reduce learning time by one-third to one-half (Regian et al., 1996). Though a through meta-analysis is wanting in the literature, certain ITSs have been shown to be 1 standard deviation better than classroom instruction(Anderson et al., 1995; Koedinger et al 1997) and have been estimated to be 1.75 standard deviations better than self-study (Corbett, 2001).

One study that was conducted in an authentic high school environment over several semesters compared the achievement of students who used an ITS (the Pump Algebra Tutor, described above) and those in traditional algebra classrooms (Koedinger et al., 1997). The students using the ITS showed dramatic gains, including 15%-25% better performance on skill knowledge and 50% to 100% better performance in problem solving than the control classroom condition. The authors claim that use of their tutor accounted for a one letter-grade improvement among students in the study (Koedinger et al, 1997, Anderson et al. 1995).

Further, because ITS developers have a vested interest in seeing their design's succeed, third-party evaluation's of tutors have been conducted. One such objective evaluation has indicated that ITSs with certain features that support affective and motivational variables lead to an increase in student motivation and measurably transformed classroom culture (Schofield, 1995). Likewise, similar studies report that ITSs teach to think more deeply about complex skills, develop enhanced reasoning, and acquire better comprehension and design skills (Roschelle et al., 2000).

All of this is not to uncritically lionize ITSs; there are still many limitations to their efficacy and applicability. Chief among these is the unfortunate fact that many developers of ITS are only superficially acquainted with research in the fields of education and psychology, and often incorrectly apply what little they do know of these theories. For instance, even a cursory search of the recent literature returns disconcertingly many studies based on student "learning styles". Aside from this gulf between learning researchers and computer scientists, there are plenty of issues characteristic of the technology itself. Most current ITSs require step-by-step interpretation of each subject in order to accurately model student ability, which is feasible only for very simple, procedure-based domains like high school mathematics.

Another pitfall inherent to the software itself are the many constraints imposed on students in order to accomplish the necessary monitoring and remediation. These environments necessarily restrict student input to terms which the program can understand, and often depend on built in expectations about the course of problem solving, limiting a learner's freedom. However, letting students pursue multiple solution strategies within a given problem may not be important for learning in certain ITSs. Wallkens et al. (2013) found that the amount of freedom offered by their system did not affect students' learning outcomes or intrinsic motivation. Still, a lock-step approach means that there is also the risk that student errors may be monitored and redressed too quickly, not giving them adequate time to learn from their mistakes.

Finally and perhaps most crucially, the instructional potential of these systems has been limited by the rigidity of teaching style with a given tutor; though numerous tutors have been designed based on as many different teaching/learning strategies, and though they adapt to student variables at many different levels, none of them as yet have demonstrated the ability to effectively accommodate students at the level of instruction for a given domain. This is perhaps due to the immense cost, in both time and money, associated with developing a working ITS based upon a single instructional technique. To be able to shift between teaching strategies would require the coordinate of several such programs within a program. However, having different tutors specialized for different subjects may not be such an issue, especially in cases where the instructional strategy is highly specific to the content itself.

5 Future Directions of ITSs

In some areas, such as reasoning about a student's knowledge and inferring the nature of student misconceptions, ITSs are alreadying performing just as well as human tutors.

In other areas, such as employing natural language dialogue or flexibly applying multiple pedagogical interventions, they are not as good. Much of what I have reported on here represents the vanguard of research and development in the field... it is important to bear in mind that current computer education used in school settings is not like this, not customized or adaptive, not social, and not integrated into the curriculum. The customization of teaching and the democratizing of its quality is perhaps the greatest potential of educational technology; today, in the developing world nearly 2 billion children receive little to no formal education (Woolf, 2009).

Several other, more long-term goals of ITS are the design of tutors that help students develop metacognitive and self-regulatory skills, such as self-awareness for dealing with failure and frustration. These skills can help students persevere through difficulties and improve their self-efficacy and motivation. Another related goal is the capacity of such a system to recognize students' emotion, and once recognized, to provide appropriate interventions that calm and encourage them, providing support for positive attitudes that outweighs or distracts from unpleasant aspects of failure. For example, if student frustration has been detected, tutors must be able to determine whether to intervene or let the "desirable difficulties" of inquiry learning take their course (Burleson and Picard, 2004). Although there is lots of current interest indeveloping "natural language" tutors that can interpret students' speech and respond in kind, a recent study foud no different in learning gains for students using spoken language versus students using text-based responses in an ITS (D'Mello et al., 2011)

Anothing interesting implication of ITSs is how to go about the grading and assessment within such a system. Should the individual be assessed even though the tutor is providing guidance throughout the process? Trivedi et al. (2011) show that assessment quality determined by assitance data from an ITS is a better estimator of student knowledge than their scores on an exam. One implication of this study is that a lot of time spent on assessment could be saved for instruction if assessment and instruction could be integrated in this fashion. The widespread adoption of ITSs is clearly going to impact education at its most fundamental levels; the roles of students and teachers, the way students are assessed, the traditional spatial/temporal classroom, the structure of curricula... the list goes on. But this is not the first time a technology has deeply impacted the way human beings learn:

> "... consider again the example of books: they have certainly outperformed people in the precision and permanence of their memory, and the reliability of their patience. For this reason, they have been invaluable to humankind. Now imagine active books that can interact with the reader to communicate knowledge at the appropriate level, selectively highlighting the interconnectedness and ramifications of items, recalling relevant information, probing understanding, explaining difficult areas in more depth, skipping over seemingly known material ... [Intelligent Tutoring Systems] are indeed an attractive dream." (Wenger, 1987)

Wenger's attractive dream is well on its way to becoming a reality!

6 A Subjective Conclusion

Although I tried to be as objective and comprehensive as possible in my previous evaluation of Intelligent Tutoring Systems, I find myself wanting to play devil's advocate a bit. This is especially true in light of the countervailing evidence others have presented about the literature of disappointment encountered while researching Computer-Based Instruction as applied to certain aspects of education. I feel like I could benefit by making more of an effort to see *the other side* of educational technology; extreme opinions and side-taking on these issues are perilous, and the most well-reasoned, even-tempered evaluation is best. However, in seeking out opinions from the other side, I discovered some that I find to be so unproductive and misguided that I will probably overreact to them here in the last analysis. This is my caveat lector.

Educators, administrators, and researchers have put a lot of stock in personlized learning, and with the promise of "a teacher for every student" (Woolf 2008, p. 15) it is easy to see why. As noted in my last presentation, "Personalized Learning" is one of the National Academy of Engineering's 14 "Grand Challenges," alongside goals like "Prevent Nuclear Terror" and "Engineer Better Medicines." But what would the implications of such an educational overhaul be? This is something that I have failed to fully consider in my previous research on this topic, and I would like to spend some time doing so in this last paper. This is *very* forward looking, and it assumes that the technology is universally embraced, but these things need to be considered. Would they really reduce costs? Would they really improve the productivity of our education system? Would the reductions in classroom activity *incidental to programmed education*, such as the social growth that comes from peer interactions or the ethics that are inclucated along the way, be detrimental to society? In short, would considerations of educational productivity, efficiency, and efficacy be strong enough to justify widespread adoption of Intelligent Tutoring Systems if it comes at the cost of certain humanistic and social considerations? Does private industry stand to gain at the expense of the nation's teachers and students?

In 1950s, B. F. Skinner, a personal hero of mine, gave the world one of the first "teaching machines," paving the way for their later use in classrooms. They were rudimentary contraptions to be sure, but they succeeded in providing immediate reinforcement to students' work and produced promising increases in test scores. In 1960, an 8^{th} grade class in Virginia went through all of 9^{th} grade algebra in a year using these primitive teaching machines; what's more, when tested the following year, their retention was above normal (Skinner 1961, p. 387; Skinner 1989). They also generated crude reports of student performance that could give a teacher insight into a specific student's difficulties. The most crucial aspect of these teaching machines, though, was the fact that students paid attention and were *then rewarded*; this is in contrast to, say, and educational video, where the reward (the entertainment and sensory stimulation) was not at all contingent on the students' attending well to the material. In 1989, in the "Letters" section of the journal Science, Skinner wrote that "the teaching profession must turn, as all other professions have turned, to instrumentation. That will not dehumanize teaching; it will free it from what is now essentially the inhumane, punitive formula of 'study and learn or else'" (Skinner, 1989).

Since the advent of these techniques, however, they have had many outspoken critics. One such critic, whose discussion inspired some of the more reactionary points that I will be making in this exposition, is a Canadian adjunct professor of education named Phil McRae. He drew a lot of attention earlier this year with a "Blog Post turned Magazine Article turned National OpEd" where he gives a very one-sided critique of adaptive learning systems, peppered liberally with rhetorical flourish but wanting sorely in substance (McCrae, 2013). I will certainly give credit where it it due him, but I want to take issue with certain of his remarks in the following few paragraphs and in so doing bring my conclusions about the future of ITS to the fore.

McCrae implies in his piece that that computer-mediated instruction will threaten family life as we know it and deprive children of their fun-loving free time, because increasingly over-involved parents will have yet another tool in their toolbox with which to separate kids from unproductive activities. On the contrary! These programs have been shown to reduce learning time while producing the same (and often better) retention and transfer (VanLehn, 2011; Koedinger et al 2009). Indeed, they would actually serve to free up time that students would've been investing in absent-minded study and busy-work boredom. To be sure, ITS have many shortcomings, but increasing the time it takes to learn is assuredly not among them. This is not to deny that something of a tutoring obsession has taken parts of the developed world by storm: for instance, in 2010 74% of all South Korean students were involved in private after-school instruction of one kind or another, and at an average cost of $2,600 per student for the year (Ripley, 2011). McCrae sees in this a slippery slope, where parents bring tutoring permanently into the home with ITSs, and makes it out to be the fault of the technology, and not the symptom of a larger societal issue which it undoubtedly is.

He also argues verbatim that "there are no quick fixes to learning and teaching" and implies that student progress takes exactly the amount of time and practice it always has (Mcrae, 2013). This attitude is lazy and irresponsible, and I cannot believe it is being trotted out by a professor of education! It is tantamount to sitting on one's hands and whole-heartedly embracing the status quo; it undermines entirely the fundamental motivations for educational research! He is dismissing out of hand what are demonstrably effective educational interventions, especially in areas like mathematics where the U.S. lags behind; these interventions have been shown unequivocally to produce positive learning outcomes! The U.S. government's educational research arm (the Institution of Education Science) maintains the "What Works Clearinghouse", an initiative that identifies legitimately effective educational interventions to encourage informed decision-making among districts nationwide. At last check, they have ranked an ITS as 5^{th} most effective out of all 37 mathematics interventions for which they have found creditable and reliable evidence of effectiveness. And it's not like this is an isolated case; another such system is similarly ranked for effectiveness in general science education. These findings are of undeniable importance and to refuse to consider them because of ideological or political concerns is, to me, unconscionable.

Further, in his article, lots of scaremongering is leveled at the teachers themselves, even though it is an unsupported reach to claim any privileged knowledge that an ITS necessarily threatens the fundamental role of the teacher in the classroom. Good teachers are the exemplars of humanity, each of whom grows through experience to become intimately aware of their students individually, their facets and uniquenesses, and knows how best to motivate and engage them. More generally, they have a complex understanding issues like inequality and poverty in the classroom and a knowledge of community, peer, and even family influences that may have important consequences for learning in the classroom. It is ridiculous to envision a computer replacing the teacher, but this is just what detractors of ITS would have you believe; they set up a straw-man by depicting ITS usage as a solitary child sitting "in front of a computer screen for hours on end" (McCrae, 2013). While the teacher's role may indeed be redefined in the future (what *won't be redefined* in the future?), it will not be anything these dramatic distopian out-of-work-overnight scenarios conjured up by hidebound traditionalists.

Nor will they entail a "narrowing of cognition." Learning is most successful when it occurs in active processes that involve social, emotional, and intrapersonal experiences; this much is almost universally accepted by educational researchers (Schallert and Martin, 2003), and it is yet another straw-man argument to discount educational technology as a stone-age stimulus-response model that denies the many irreplaceable aspects of learning alongside (and directly from) your peers in a classroom. Indeed, learning to deal with the uncertainties and anxieties of classroom life teach students one of the most important skills in today's rapidly changing world: the ability to make sound judgments under uncertainty. Shulman (2005) argues that "an absence of emotional investment, even risk and fear, leads to an absence of intellectual and formational yield," and that the social anxieties of the classroom are necessary features for durable learning. No one is suggesting that these considerations don't matter, or that computers could provide viable substitutes for these fundamental aspects of schooling. Instead, what ITS claims to offer for education are exactly the kinds of things that machines do best; ITSs act as a supplement to traditional curricula in very well-defined domains like math, science, and computer programming, because they are fine-tuned, pedagogically consistent, and probabilistically-adaptive to the needs demonstrated by individual students.

Mechanizing even a few minor details of the *existing educational paradigm* can also have powerful effects. To take a simple example, the now-critical, very visible, and seemingly insoluble issue of growing class-size could be rendered far less severe with technologies like these in the classroom. Another practical, even simpler use for these technologies is in student evaluation. Multiple choice testing isn't likely to go anywhere because it is far more convenient and efficient than other methods of assessment, and provided they contain high-quality items and test valid concepts, at least as effective. They are also more objective than other forms of assessment, because there is no teacher in the loop to interpret answers in a biased way, consciously or unconsciously. They are often criticized, though, for producing only superficial learning, and providing assessment through recognition. In the hands of an ITS, however, multiple choice testing could really be improved upon. Roediger and Butler (2010) note that, in addition to the benefits of retrieval practice during testing, feedback has particularly important implications for multiple choice testing. If feedback is provided after a multiple choice test, students learn more from the test, and any negative effects (like potentially encoding the "trap answers") are nullified. They report on one study that administered a multiple-choice pretest to students before a criterion test; one version provided feedback immediately after making a response, another version provided no feedback, and another group, the control, did not take the multiple-choice pretest. Simplying taking the pretest was extremely beneficial for recall, but when correct-answer feedback was given after each item, performance increased another 10%. As an added bonus, research suggests that computerized feedback can be even more effective than non-computerized feedback (Baumeister et al, 1990). Areas like this, where a machine no more elaborate than Skinner's "teaching machine" produces improvements upon already entrenched educational practice, point to the great untapped potential for this technology in education.

There are certainly a few areas where we find reasonable reservations about embracing educational technologies like ITS. One such area where critics have made a cogent counter to ITS is in the creative and the performing arts; no computer program, no matter how advanced, could foster creativity in the arts like a human teacher or mentor could. But no one is saying otherwise! Another criticism that really holds water is the threat of private companies profiteering off of student-generated data by setting up incentives for the production of such data (i.e., their brand of learning surveillance system, masquerading as an effective educational intervention). This is an absolutely valid concern, and one I think we would be wise to be vigilant about. Public schools should guard students' personal data against those who would stand to gain from it. But in no way should this threat be seen as a wholesale condemnation of educational technology, especially those that have shown promise in ameliorating areas of weakness in our education system. It's a "baby out with the bathwater" reaction, and one at cross-purposes with educational progress.

On a very personal and therefore completely subjective note, I feel like I am the kind of person who would respond quite well to instruction provided by an Intelligent Tutoring System. When I was a Junior in high school, my physics class was required to submit their homework online (through some system that was actually connected with the University of Texas, if I remember correctly). You would work your problem, and then enter it in, and you could immediately see whether you were right or wrong; also, it would provide the solution steps so you wouldn't systematically botch the rest of the assignment by repeating your mistake. I loved this. I felt academically liberated from the control of the teacher, of whom I wasn't entirely fond. This way, it kept her entirely out of the loop; I didn't have to perform for her or for anyone else; my work was between me and an unjudgmental, disinterested website. I respond very negatively to perceived control of my behavior by others, and it is especially bad to work to appease someone you find deeply offensive. Indeed, there was a point in high school that I stopped doing my work all together because I felt like my mother was taking the credit for my achievements (most of which, I now concede, she was probably entitled to). Petty, yes, but Skinner tells us that these feelings are endemic, particularly in Western societies: "the struggle for freedom is concenred with stimuli intentionally arranged by other people" (Skinner, 1971, p. 39). When the external contingencies the control one's behavior are obvious and unfavorable to that person, "the person cannot then be said to be free, and the contingencies deserve the credit" (Skinner, 1971, p. 77). Sorry to be so Skinner-y in this paper! I just read Beyond Freedom and Dignity for the first time (During finals, I know! What else can one do while trapped on the bus for an hour and a half each day?), and so the ideas are fresh in my mind.

To conclude, I find that after much critical exploration of Intelligent Tutoring Systems I still endorse many of the beliefs I naively esposed as my interest in this topic was newly burgeoning. I still believe that, in the hands of caring teachers and within the confines of traditional classrooms, ITS can provide a effective, engaging supplement to certain kinds of subject-matter instruction with all the advantages of individual tutoring. I believe it can do this without requiring any drastic overhaul of education, any indignities to the role of the teacher, or any prohibitive costs. I was struck by the seemingly limitless educational potential of these tutors: they can implement mastery learning, scaffolding understanding, facilitate self-regulation, provide authentic learning contexts, account for and respond to moods, motivation, and interest, and adapt a given instance of instruction to the specific student based on methods that proved effective in past interactions with other students having similar difficulties (Woolf, 2008). Though it would be foolhardy put too much stock in technology where education is concerned, I think that Intelligent Tutoring Systems have a place in instruction and, with time, will bring us much closer to optimizing learning for all.

Baumeister, R. F., Hutton, D. G., Cairns, K. J. (1990). Negative Effects of Praise on Skilled Performance. Basic Applied Social Psychology, 11(2), 131–148.

Koedinger, K. R., & Aleven, V. (2007). Exploring the assistance dilemma in experiments with cognitive tutors. Educational Psychology Review, 19(3), 239–264. doi:10.1007/s10648-007-9049-0

McRae, P. (2013). Rebirth of the Teaching Machine Through the Seduction of Data Analytics. ATA Magazine, 93(4), 30–36.

Ripley, A. (2011, September 25). Teacher, Leave Those Kids Alone. Time. Retrieved from http://content.time.com/time/magazine/article/0,9171,2094427,00.html

Roediger III, H. L., Butler, A. C. (2011). The critical role of retrieval practice in long-term retention. Trends in Cognitive Sciences, 15(1), 20–27. doi:10.1016/j.tics.2010.09.003

Shulman, L. S. (2005). Signature pedagogies in the professions. Daedalus, 134(3), 52–59. doi:10.1162/0011526054622015

Skinner, B. F. (1961). Why We Need Teaching Machines. Harvard Educational Review, 31(4), 377–398.

Skinner, B .F. (1989). From Behaviorism to Teaching Machines to Enjoying Old-Age - a Citation Classic Commentary on the Behavior of Organisms - an Experimental-Analysis by Skinner,b.f. Current Contents/Clinical Medicine, (6), 16–16.

Skinner, B. F. (1971). Beyond Freedom and Dignity. Hackett Publishing.

VanLehn, K. (2011). The Relative Effectiveness of Human Tutoring, Intelligent Tutoring Systems, and Other Tutoring Systems. Educational Psychologist, 46(4), 197–221. doi:10.1080/00461520.2011.611369

Annotated Bibliography

Though in my previous paper I included almost 100 sources, most of these were used in a very cursory way, and many of them I found because they were sited by my more primary sources. This list contains a few of the more influential papers for me that I consulted for this project.

Woolf, B. P. (2009). Building intelligent interactive tutors: student-centered strategies for revolutionizing e-learning. Amsterdam; Boston: Morgan Kaufmann Publishers/Elsevier.

I read this entire 400 page volume over the course of the semester, and it directed my foray into the Intelligent Tutoring Systems literature more than any other source. Beverly Woolf Ed.D., Ph.D. is a professor of computer science at Amherst whose entire career has centered on computer-based instructional technology. In this book, she gives charts the development of these technologies overtime framed in a perspective that unifies approaches from artificial intelligence with educational research. It is organized into three parts: the first is an introduction to these technologies that addresses certain issues and features, setting the stage for Part 2. In the second part, especially in chapter 4 on "Teaching Knowledge," she shows how these tutors have been effectively implemented in areas such as modeling student knowledge and affect, how they have been programmed with instructional techniques that have support in the learning literature (Chapter 4, Section 3), and how they have been enriched with technological innovations such as immersive interfaces, natural language processing, and more authentic communication. In part 3, she addresses the more technical aspects of how to go about programming these tutors to achieve these complex ends (Chapter 7). She also discusses some newly emerging "collaborative inquiry" tutors that try to break the mold by supporting more situated and constructivist approaches (Chapter 8), and provides of survey of the state-of-the-art in web-based learning environments (Chapter 9).

VanLEHN, K. (2011). The Relative Effectiveness of Human Tutoring, Intelligent Tutoring Systems, and Other Tutoring Systems. Educational Psychologist, 46(4), 197–221. doi:10.1080/00461520.2011.611369

This paper is one of the only reputable meta-analyses of ITS that I discovered in my research. He reviews studies that have compared human tutors to ITSs, and that have a "no tutoring" control. This paper provided my sources for effect sizes, where were found to be d=0.79 above the control condition for human tutoring and 0.76 above controls for ITS. He also talks about his "granularity hypothesis" for why previous incarnations of computer-based instruction produces weaker effects than their "intelligent" counterparts, which break each problem down into subparts, further reducing those subparts into substeps, so they can really get at the issue a student is encountering with the material.

ALEVEN, V., ROLL, I., McLAREN, B. M., & KOEDINGER, K. R. (2010). Automated, Unobtrusive, Action-by-Action Assessment of Self-Regulation During Learning With an Intelligent Tutoring System. Educational Psychologist, 45(4), 224–233. doi:10.1080/00461520.2010.517740

This article reviews several features of a Cognitive Tutors meant to train four different kinds of self-regulatory skills: help-seeking, self-explanation, gaming the system, and error self-correction. Metacognition is notoriously difficult to measure, and because these methods are unobtrusive, they do not distract from instructional goals, can be applied longitudinally, and are not subject to distortion through student self-report. The same is not true for traditional measures of self-regulated learning (think aloud approaches, or film-prompted interviewing), which are subject to interference by feelings of knowing and judgments of learning that may not be accurate. The main takeaway from this review is that ITSs are excellent at producing event-based data regarding student learning, and the same holds true for student self-regulation.

Koedinger, K. R., 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 (IJAIED), 8, 30–43.

This is another paper I relied upon for proof-of-concept material; it demonstrates the effectiveness of Carnegie Learning's Cognitive Tutor (specifically, the Pump Algebra Tutor) in high school classrooms in a large-scale experiment in the 1993-1994 school year. These researchers built this algebra tutor to support a mathematics curriculum developed by the Pittsburgh Urban Mathematics Project, and it was used by 470 students. Students who used this tutor as part of the curriculum saw average increases of 15 % on standardized tests. What I liked most about this study was that it illustrated how ITS can work "in the wild" in urban Pittsburgh high schools.

Anderson, J. R., Corbett, A. T., Koedinger, K. R., & Pelletier, R. (1995). Cognitive tutors: Lessons learned. Journal of the Learning Sciences, 4(2), 167–207

Along with Berverly Woolf's book, this was one of my original sources; it was also the paper that introduced me to John Anderson and colleagues, as well as ACT-R theory. It provides a history of the development of theirearly LISP, geometry, and algebra tutors, and they reported on the significant gains that even these incipient efforts produced. Immediate feedback was crucial to achieving these outcomes, the most effective tutorial interaction style was as a tool to be used for learning (not as a digital humanoid agent that interacts in contrived ways, a seductive but ineffective pitfall for many tutor developers). Anderson also discusses transfering knowledge gained within the tutor to new environments, and how this can be achieved through a problem-solving environment. Though the paper is almost 20 years old, it was a good source for some of my historical material and a foundational source for my understanding of CMU-style cognitive tutors.

Pavlik, P.,& Toth, J. (2010). How to Build Bridges between Intelligent Tutoring System Subfields of Research. In V. Aleven, J. Kay, & J. Mostow (Eds.), Intelligent Tutoring Systems, Part Ii (Vol. 6095, pp. 103–112). Berlin: Springer-Verlag Berlin.

What I liked most about this paper was that it recognized the interdisciplinary nature of Intelligent Tutoring Systems; they require efforts from various fields: artificial intelligence, statistics, cognitive psychology and educational psychology, human-computer interaction, the list goes on and on. I like how ITSs represent a common goal that unites this disparate disciplines and their different perspectives in harmony with a common purpose. The richness of these different perspectives is useful in finding what is effective, and it allows for many new combinations of ideas. When these disciplines cannot see eye to eye, though, these authors suggest ways of "building bridges".

References

Advances in intelligent tutoring systems (No. volume 308). (2010). Berlin: Springer.

Alan Lesgold, S. L. (1992). SHERLOCK: a coached practice environment for an electronics troubleshooting job.

Aleven, V., & Koedinger, K. R. (2002, April). An effective metacognitive strategy: learning by doing and explaining with a computer-based cognitive tutor. *Cognitive Science*, *26*(2), 147–179. (WOS:000175489400001) doi: 10.1016/S0364-0213(02)00061-7

ALEVEN, V., ROLL, I., McLAREN, B. M., & KOEDINGER, K. R. (2010). Automated, unobtrusive, action-by-action assessment of self-regulation during learning with an intelligent tutoring system. *Educational Psychologist*, *45*(4), 224–233. Retrieved 2013-10-21, from `http://www.tandfonline.com/doi/abs/10.1080/00461520.2010.517740` doi: 10.1080/00461520.2010.517740

Ally, M. (2005). Designing effective computer-based learning materials. In C. Howard, J. V. Boettcher, & L. Justice (Eds.), *Encyclopedia of distance learning* (Vol. 2, pp. 525–533). Hershey, PA: Idea Group Reference. Retrieved 2013-09-23, from

Anderson, J. (1982). Acquisition of cognitive skill. *Psychological Review*, *89*(4), 369–406. (WOS:A1982NY77100003) doi: 10.1037//0033-295X.89.4.369

Anderson, J. (1993, January). Problem-solving and learning. *American Psychologist*, *48*(1), 35–44. (WOS:A1993KJ77400004) doi: 10.1037//0003-066X.48.1.35

Anderson, J., Conrad, F., & Corbett, A. (1989, December). Skill acquisition and the lisp tutor. *Cognitive Science*, *13*(4), 467–505. (WOS:A1989CH32100001) doi: 10.1016/0364-0213(89)90021-9

Anderson, J. R., Corbett, A. T., Koedinger, K. R., & Pelletier, R. (1995). Cognitive tutors: Lessons learned. *Journal of the Learning Sciences*, *4*(2), 167–207. (WOS:A1995TQ89500002) doi: 10.1207/s15327809jls0402$_2$

Antal, M., & Koncz, S. (2011, June). Student modeling for a web-based self-assessment system. *Expert Systems with Applications*, *38*(6), 6492–6497. Retrieved 2013-09-12, from `http://www.sciencedirect.com/science/article/pii/S0957417410013345` doi: 10.1016/j.eswa.2010.11.096

Arroyo, I., Beal, C., Murray, T., Walles, R., & Woolf, B. (n.d.). *Wayang outpost: Intelligent tutoring for high stakes achievement tests.*

Arroyo, I., Murray, T., Beck, J. E., Woolf, B. P., & Beal, C. R. (2003). Formative evaluation of AnimalWatch. In H. U. Hoppe, F. Verdejo, & J. Kay (Eds.), *Artificial intelligence in education: Shaping the future of learning through intelligent technologies* (Vol. 97, pp. 371–373). Amsterdam: I O S Press. (WOS:000189487400046)

Artificial intelligence and tutoring systems: Computational and cognitive approaches to the communication of knowledge: E wenger, morgan kaufman inc (1987) 486pp £29.95. (1988a, December). *Knowledge-Based Systems*, *1*(5), 311. Retrieved 2013-09-13, from http://www.sciencedirect.com/science/article/pii/0950705188900913 doi: 10.1016/0950-7051(88)90091-3

Artificial intelligence and tutoring systems: Computational and cognitive approaches to the communication of knowledge: E wenger, morgan kaufman inc (1987) 486pp £29.95. (1988b, December). *Knowledge-Based Systems*, *1*(5), 311. Retrieved 2013-11-13, from http://www.sciencedirect.com/science/article/pii/0950705188900913 doi: 10.1016/0950-7051(88)90091-3

Baker, R. S. J. D., D'Mello, S. K., Rodrigo, M. M. T., & Graesser, A. C. (2010, April). Better to be frustrated than bored: The incidence, persistence, and impact of learners' cognitive affective states during interactions with three different computer-based learning environments. *International Journal of Human-Computer Studies*, *68*(4), 223–241. (WOS:000276129500004) doi: 10.1016/j.ijhcs.2009.12.003

Bangert-Drowns, R. L., Kulik, C.-L. C., Kulik, J. A., & Morgan, M. (1991, July). The instructional effect of feedback in test-like events. *Review of Educational Research*, *61*(2), 213–238. Retrieved 2013-11-13, from http://www.jstor.org/stable/1170535 (ArticleType: research-article / Full publication date: Summer, 1991 / Copyright © 1991 American Educational Research Association) doi: 10.3102/00346543061002213

Baumeister, R. F., Hutton, D. G., & Cairns, K. J. (1990, June). Negative effects of praise on skilled performance. *Basic & Applied Social Psychology*, *11*(2), 131–148. Retrieved 2013-11-13, from http://ezproxy.lib.utexas.edu/login?url=http://search.ebscohost.com/login.aspx?direct=truedb=a9hAN=7301056site=ehost-li

Bickmore, T. W., Vardoulakis, L. M. P., & Schulman, D. (2013, September). Tinker: a relational agent museum guide. *Autonomous Agents and Multi-Agent Systems*, *27*(2), 254–276. (WOS:000318297400005) doi: 10.1007/s10458-012-9216-7

Boulay, B. D., & Luckin, R. (2001). Modelling human teaching tactics and strategies for tutoring systems. *International Journal of Artificial Intelligence in Education*, *12*, 235–256.

Brown, J. S., Collins, A., & Duguid, P. (1989, January). Situated cognition and the culture of learning. *Educational Researcher*, *18*(1), 32–42. Retrieved 2013-11-13, from http://edr.sagepub.com/content/18/1/32 doi: 10.3102/0013189X018001032

Burleson, W. (2006). *Affective learning companions: Strategies for empathetic agents with real-time multimodal affective sensing to foster meta-cognitive and meta-affective approaches to learning, motivation, and perseverance* (Ph.D., Massachusetts Institute of Technology, United States – Massachusetts). Retrieved 2013-11-13, from http://search.proquest.com/docview/304950932

Chang, T.-Y., & Ke, Y.-R. (2013, January). A personalized e-course composition based on a genetic algorithm with forcing legality in an adaptive learning system. *Journal of Network and Computer Applications*, *36*(1), 533–542. (WOS:000312683300048) doi: 10.1016/j.jnca.2012.04.002

Chapman, R. J., Ryder, J., Bell, B., Wischusen, D., & Benton, D. (2004, September). Strata (synthetic teammates for realtime anywhere training and assessment): An integration of cognitive models and virtual environments for mobile scenario based training. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, *48*(17), 2094–2098. Retrieved 2013-11-13, from http://pro.sagepub.com/content/48/17/2094 doi: 10.1177/154193120404801703

Cheikes, B. A., Geier, M., Hyland, R., Linton, F., Riffe, A. S., Rodi, L. L., & Schaefer, H.-P. (1999). Embedded training for complex information systems. *INTERNATIONAL JOURNAL OF ARTIFICIAL INTELLIGENCE IN EDUCATION*, *10*, 314—314. Retrieved from http://citeseerx.ist.psu.edu/viewdoc/summary?doi=10.1.1.133.6784

Chen, C.-M. (2008, September). Intelligent web-based learning system with personalized learning path guidance. *Computers & Education*, *51*(2), 787–814. (WOS:000257013800021) doi: 10.1016/j.compedu.2007.08.004

Chen, C.-M., & Duh, L.-J. (2008, May). Personalized web-based tutoring system based on fuzzy item response theory. *Expert Systems with Applications*, *34*(4), 2298–2315. (WOS:000253521900008) doi: 10.1016/j.eswa.2007.03.010

Chen, C. M., Lee, H. M., & Chen, Y. H. (2005, April). Personalized e-learning system using item response theory. *Computers & Education*, *44*(3), 237–255. (WOS:000226574400003) doi: 10.1016/j.compedu.2004.01.006

Chou, C. Y., Chan, T. W., & Lin, C. J. (2003, April). Redefining the learning companion: the past, present, and future of educational agents. *Computers & Education*, *40*(3), 255–269. (WOS:000181047800004) doi: 10.1016/S0360-1315(02)00130-6

Chou, C.-Y., Huang, B.-H., & Lin, C.-J. (2011, December). Complementary machine intelligence and human intelligence in virtual teaching assistant for tutoring program tracing. *Computers & Education*, *57*(4), 2303–2312. (WOS:000295860100011) doi: 10.1016/j.compedu.2011.06.005

Chrysafiadi, K., & Virvou, M. (2013, September). Student modeling approaches: A literature review for the last decade. *Expert Systems with Applications*, *40*(11), 4715–4729. Retrieved 2013-09-12, from http://www.sciencedirect.com/science/article/pii/S095741741300122X doi: 10.1016/j.eswa.2013.02.007

Colace, F., & De Santo, M. (2010). Ontology for e-learning: A bayesian approach. *IEEE Transactions on Education*, *53*(2), 223–233. doi: 10.1109/TE.2009.2012537

Conati, C., Gertner, A., & Vanlehn, K. (2002, November). Using bayesian networks to manage uncertainty in student modeling. *User Modeling and User-Adapted Interaction*, *12*(4), 371–417. (WOS:000179499600002) doi: 10.1023/A:1021258506583

Corbett, A., & Anderson, J. (2008, November). Student modeling and mastery learning in a computer-based programming tutor. *Department of Psychology*. Retrieved from http://repository.cmu.edu/psychology/18

Corbett, A., McLaughlin, M., & Scarpinatto, K. C. (2000). Modeling student knowledge: Cognitive tutors in high school and college. *User Modeling and User-Adapted Interaction*, *10*(2-3), 81–108. (WOS:000166528800002) doi: 10.1023/A:1026505626690

Corbett, A. T. (1995). *Knowledge tracing: modeling the acquisition of procedural knowledge* (No. 95-129). Pittsburgh, Pa: School of Computer Science, Carnegie Mellon University.

Creating an adaptive training system: Integration of the SMART student model in a RIDES tutor - FS97-01-005.pdf. (n.d.). Retrieved 2013-09-25, from http://aaaipress.org/Papers/Symposia/Fall/1997/FS-97-01/FS97-01-005.pdf

Data mining in e-learning (No. v. 4). (2006). Southampton: WIT.

Desiano, S. D. (2002). Expert systems. In R. R. Flynn (Ed.), *Computer sciences* (Vol. 3, pp. 87–91). New York: Macmillan Reference USA. Retrieved 2013-09-23, from

D'mello, S., & Graesser, A. (2013, January). AutoTutor and affective autotutor: Learning by talking with cognitively and emotionally intelligent computers that talk back. *ACM Trans. Interact. Intell. Syst.*, *2*(4), 23:1–23:39. Retrieved 2013-09-25, from http://doi.acm.org.ezproxy.lib.utexas.edu/10.1145/2395123.2395128 doi: 10.1145/2395123.2395128

D'Mello, S., Olney, A., Williams, C., & Hays, P. (2012, May). Gaze tutor: A gaze-reactive intelligent tutoring system. *International Journal of Human-Computer Studies*, *70*(5), 377–398. (WOS:000303303900005) doi: 10.1016/j.ijhcs.2012.01.004

D'Mello, S. K., Craig, S. D., Witherspoon, A., McDaniel, B., & Graesser, A. (2008, February). Automatic detection of learner's affect from conversational cues. *User Modeling and User-Adapted Interaction*, *18*(1-2), 45–80. (WOS:000252678000003) doi: 10.1007/s11257-007-9037-6

D'Mello, S. K., Dowell, N., & Graesser, A. (2011, March). Does it really matter whether students' contributions are spoken versus typed in an intelligent tutoring system with natural language? *Journal of Experimental Psychology-Applied*, *17*(1), 1–17. (WOS:000288776000001) doi: 10.1037/a0022674

Drummond, J., & Litman, D. (2010, January). In the zone: Towards detecting student zoning out using supervised machine learning. In V. Aleven, J. Kay, & J. Mostow (Eds.), *Intelligent tutoring systems* (pp. 306–308). Springer Berlin Heidelberg. Retrieved 2013-09-23, from http://link.springer.com.ezproxy.lib.utexas.edu/chapter/10.1007/978-3-642-13437-1$_5$3

Dweck, C. S. (1986). Motivational processes affecting learning. *American Psychologist*, *41*(10), 1040–1048. doi: 10.1037/0003-066X.41.10.1040

Ebbinghaus, H. (1913). *Memory; a contribution to experimental psychology* (No. no. 3). New York city: Teachers college, Columbia university.

Elliot, A. J., & Church, M. A. (1997, January). A hierarchical model of approach and avoidance achievement motivation. *Journal of Personality and Social Psychology*, *72*(1), 218–232. (WOS:A1997WC93600019) doi: 10.1037//0022-3514.72.1.218

Feldon, D. F. (2007). Cognitive load and classroom teaching: The double-edged sword of automaticity. *Educational Psychologist*, *42*(3), 123–137. (WOS:000249221000001)

Fox, J.-P. (2010). *Bayesian item response modeling.* Dordrecht: Springer.

García, P., Amandi, A., Schiaffino, S., & Campo, M. (2007, November). Evaluating bayesian networks' precision for detecting students' learning styles. *Computers & Education*, *49*(3), 794–808. Retrieved 2013-09-12, from `http://www.sciencedirect.com/science/article/pii/S0360131505001788` doi: 10.1016/j.compedu.2005.11.017

Gertner, A. S., & VanLehn, K. (2000a). Andes: A coached problem solving environment for physics. In G. Gauthier, C. Frasson, & K. VanLehn (Eds.), *Intelligent tutoring systems, proceedings* (Vol. 1839, pp. 133–142). Berlin: Springer-Verlag Berlin. (WOS:000171336100017)

Gertner, A. S., & VanLehn, K. (2000b, January). Andes: A coached problem solving environment for physics. In G. Gauthier, C. Frasson, & K. VanLehn (Eds.), *Intelligent tutoring systems* (pp. 133–142). Springer Berlin Heidelberg. Retrieved 2013-11-13, from `http://link.springer.com/chapter/10.1007/3-540-45108-0`$_1$7

Goleman, D. (1996). *Emotional intelligence: Why it can matter more than IQ.* Bloomsbury.

Graesser, A. C., Chipman, P., Haynes, B. C., & Olney, A. (2005, November). AutoTutor: an intelligent tutoring system with mixed-initiative dialogue. *Ieee Transactions on Education*, *48*(4), 612–618. (WOS:000233264500008) doi: 10.1109/TE.2005.856149

Graesser, A. C., VanLehn, K., Rose, C. P., Jordan, P. W., & Harter, D. (2001). Intelligent tutoring systems with conversational dialogue. *Ai Magazine*, *22*(4), 39–51. (WOS:000180288000004)

Hamdi, M. S. (2005). Support of online learning through intelligent programs. In C. Howard, J. V. Boettcher, & L. Justice (Eds.), *Encyclopedia of distance learning* (Vol. 4, pp. 1716–1724). Hershey, PA: Idea Group Reference. Retrieved 2013-09-23, from

Hausmann, R. G. M., & VanLehn, K. (2010, January). The effect of self-explaining on robust learning. *International Journal of Artificial Intelligence in Education*, *20*(4), 303–332. Retrieved 2013-11-13, from `http://dx.doi.org/10.3233/JAI-2010-010` doi: 10.3233/JAI-2010-010

How people learn: Brain, mind, experience, and school: Expanded edition. (n.d.). Retrieved 2013-09-25, from `http://www.nap.edu/openbook.php?record`$_id = 9853page = 155$

Hsu, C.-K., Hwang, G.-J., & Chang, C.-K. (2010, August). Development of a reading material recommendation system based on a knowledge engineering approach. *Computers & Education*, *55*(1), 76–83. (WOS:000277110400008) doi: 10.1016/j.compedu.2009.12.004

Hua, S. (2004). *BITS: a bayesian intelligent tutoring system for computer programming* (M.Sc., The University of Regina (Canada), Canada). Retrieved 2013-10-20, from `http://search.proquest.com.ezproxy.lib.utexas.edu/docview/305169079/abstract?accountid=7118`

Huang, M.-J., Chiang, H.-K., Wu, P.-F., & Hsieh, Y.-J. (2013). A multi-strategy machine learning student modeling for intelligent tutoring systems based on blackboard approach. *Library Hi Tech*, *31*(2), 274–293. (WOS:000322803100006) doi: 10.1108/07378831311329059

Hwang, G. J. (2003, April). A conceptual map model for developing intelligent tutoring systems. *Computers & Education*, *40*(3), 217–235. (WOS:000181047800002) doi: 10.1016/S0360-1315(02)00121-5

Hwang, G.-J., Tseng, J. C. R., & Hwang, G.-H. (2008). Diagnosing student learning problems based on historical assessment records. *Innovations in Education and Teaching International*, *45*(1), 77–89. (WOS:000252459500009) doi: 10.1080/14703290701757476

Johnson, W. L., Rickel, J. W., & Lester, J. C. (2000). Animated pedagogical agents: face-to-face interaction in interactive learning environments. *INTERNATIONAL JOURNAL OF ARTIFICIAL INLIGENCE IN EDUCATION*, *11*, 47–78.

Kaklauskas, A., Vlasenko, A., Raudonis, V., Zavadskas, E. K., Gudauskas, R., Seniut, M., ... Kaklauskas, G. (2013, January). Student progress assessment with the help of an intelligent pupil analysis system. *Engineering Applications of Artificial Intelligence*, *26*(1), 35–50. (WOS:000313855000004) doi: 10.1016/j.engappai.2012.01.006

Kluger, A. N., & DeNisi, A. (1996). The effects of feedback interventions on performance: A historical review, a meta-analysis, and a preliminary feedback intervention theory. *Psychological Bulletin*, *119*(2), 254–284. doi: 10.1037/0033-2909.119.2.254

Koedinger, K. R., & Aleven, V. (2007, September). Exploring the assistance dilemma in experiments with cognitive tutors. *Educational Psychology Review*, *19*(3), 239–264. (WOS:000249403600002) doi: 10.1007/s10648-007-9049-0

Koski, T., & Noble, J. (2009). *Bayesian networks : An introduction* (2nd ed.). Hoboken: Wiley.

Kulik, C., & Kulik, J. (1991). Effectiveness of computer-based instruction - an updated analysis. *Computers in Human Behavior*, *7*(1-2), 75–94. (WOS:A1991FB47000007) doi: 10.1016/0747-5632(91)90030-5

Lave, J., & Wenger, E. (1991). *Situated learning: Legitimate peripheral participation.* Cambridge University Press.

Lepper, M. R., & Gurtner, J.-L. (1989, February). Children and computers: Approaching the twenty-first century. *American Psychologist*, *44*(2), 170–178. Retrieved 2013-11-13, from http://ezproxy.lib.utexas.edu/login?url=http://search.ebscohost.com/login.aspx?direct=truedb=pdhAN=1989-27576-001site=e doi: 10.1037/0003-066X.44.2.170

Liao, S. H. (2005, January). Expert system methodologies and applications - a decade review from 1995 to 2004. *Expert Systems with Applications*, *28*(1), 93–103. (WOS:000225261500009) doi: 10.1016/j.eswa.2004.08.003

Lin, T., & Kinshuk. (2005). Cognitive profiling in life-long learning. In C. Howard, J. V. Boettcher, & L. Justice (Eds.), *Encyclopedia of distance learning* (Vol. 1, pp. 245–255). Hershey, PA: Idea Group Reference. Retrieved 2013-09-23, from

Mathan, S. A., & Koedinger, K. R. (2005). Fostering the intelligent novice: Learning from errors with metacognitive tutoring. *Educational Psychologist*, *40*(4), 257–265. Retrieved 2013-11-13, from http://www.tandfonline.com/doi/abs/10.1207/s15326985ep4004$_7$ doi: 10.1207/s15326985ep4004$_7$

Matsuda, N., Yarzebinski, E., Keiser, V., Raizada, R., Stylianides, G. J., Cohen, W. W., & Koedinger, K. R. (2011). Learning by teaching SimStudent - an initial classroom baseline study comparing with cognitive tutor. In G. Biswas, S. Bull, J. Kay, & A. Mitrovic (Eds.), *Artificial intelligence in education* (Vol. 6738, pp. 213–221). Berlin: Springer-Verlag Berlin. (WOS:000311898000029)

McQuiggan, S. W., Mott, B. W., & Lester, J. C. (2008, February). Modeling self-efficacy in intelligent tutoring systems: An inductive approach. *User Modeling and User-Adapted Interaction*, *18*(1-2), 81–123. (WOS:000252678000004) doi: 10.1007/s11257-007-9040-y

Millán, E., Agosta, J. M., & Pérez de la Cruz, J. L. (2001). Bayesian student modeling and the problem of parameter specification. *British Journal of Educational Technology*, *32*(2), 171–181. Retrieved 2013-09-12, from http://onlinelibrary.wiley.com.ezproxy.lib.utexas.edu/doi/10.1111/1467-8535.00188/abstract doi: 10.1111/1467-8535.00188

Millán, E., Loboda, T., & Pérez-de-la Cruz, J. L. (2010, December). Bayesian networks for student model engineering. *Computers & Education*, *55*(4), 1663–1683. Retrieved 2013-09-12, from http://www.sciencedirect.com/science/article/pii/S0360131510002022 doi: 10.1016/j.compedu.2010.07.010

Moocs data offers promise of perfect teaching. (n.d.). Retrieved 2013-11-13, from http://www.bbc.com/future/story/20131029-we-can-build-the-perfect-teacher

Narciss, S. (2004). The impact of informative tutoring feedback and self-efficacy on motivation and achievement in concept learning. *Experimental Psychology*, *51*(3), 214–228. (WOS:000223329100006) doi: 10.1027/1618-3169.51.3.214

Olson, G. M., & Olson, J. S. (2003). Human-computer interaction: Psychological aspects of the human use of computing. *Annual Review of Psychology*, *54*(1), 491–516. Retrieved 2013-11-11, from http://www.annualreviews.org/doi/abs/10.1146/annurev.psych.54.101601.145044 (PMID: 12209025) doi: 10.1146/annurev.psych.54.101601.145044

Papert, S. (1993). *The children's machine: rethinking school in the age of the computer*. New York: BasicBooks.

Pavlik, P., & Toth, J. (2010). How to build bridges between intelligent tutoring system subfields of research. In V. Aleven, J. Kay, & J. Mostow (Eds.), *Intelligent tutoring systems, part ii* (Vol. 6095, pp. 103–112). Berlin: Springer-Verlag Berlin. (WOS:000280420200011)

Perkins, D. (1993). Technology meets constructivism: Do they make a marriage? In *Constructivism and the technology of instruction : a conversation* (pp. 45–54). Hillsdale N.J.: Lawrence Erlbaum Associates Publishers. Retrieved 2010-01-29, from

Roll, I., Aleven, V., McLaren, B. M., & Koedinger, K. R. (2011, April). Improving students' help-seeking skills using metacognitive feedback in an intelligent tutoring system. *Learning and Instruction*, *21*(2), 267–280. (WOS:000287121200011) doi: 10.1016/j.learninstruc.2010.07.004

Roll, I., Aleven, V., McLaren, B. M., Ryu, E., Baker, R. S. J. d., & Koedinger, K. R. (2006, January). The help tutor: Does metacognitive feedback improve students' help-seeking actions, skills and learning? In M. Ikeda, K. D. Ashley, & T.-W. Chan (Eds.), *Intelligent tutoring systems* (pp. 360–369). Springer Berlin Heidelberg. Retrieved 2013-11-13, from http://link.springer.com/chapter/10.1007/11774303$_{3}$6

Roschelle, J. M., Pea, R. D., Hoadley, C. M., Gordin, D. N., & Means, B. M. (2000, October). Changing how and what children learn in school with computer-based technologies. *The Future of Children*, *10*(2), 76–101. Retrieved 2013-11-13, from http://www.jstor.org/stable/1602690 (ArticleType: research-article / Issue Title: Children and Computer Technology / Full publication date: Autumn - Winter, 2000 / Copyright © 2000 Princeton University) doi: 10.2307/1602690

Sabo, K. E., Atkinson, R. K., Barrus, A. L., Joseph, S. S., & Perez, R. S. (2013, July). Searching for the two sigma advantage: Evaluating algebra intelligent tutors. *Computers in Human Behavior*, *29*(4), 1833–1840. (WOS:000319090200068) doi: 10.1016/j.chb.2013.03.001

Schatz, S., Oakes, C., Folsom-Kovarik, J. T., & Dolletski-Lazar, R. (2012, March). ITS + SBT: a review of operational situated tutors. *Military Psychology*, *24*(2), 166–193. Retrieved 2013-09-23, from http://ezproxy.lib.utexas.edu/login?url=http://search.ebscohost.com/login.aspx?direct=truedb=pdhAN=2012-10968-006site= doi: 10.1080/08995605.2012.672910

Schwonke, R., Ertelt, A., Otieno, C., Renkl, A., Aleven, V., & Salden, R. J. C. M. (2013, February). Metacognitive support promotes an effective use of instructional resources in intelligent tutoring. *Learning and Instruction*, *23*, 136–150. (WOS:000311770100014) doi: 10.1016/j.learninstruc.2012.08.003

Schwonke, R., Renkl, A., Krieg, C., Wittwer, J., Aleven, V., & Salden, R. (2009, March). The worked-example effect: Not an artefact of lousy control conditions. *Computers in Human Behavior*, *25*(2), 258–266. (WOS:000263779000002) doi: 10.1016/j.chb.2008.12.011

ScienceDirect (Online service). (1985). *Computers in human behavior.* Elmsford, N.Y: Pergamon Press.

Shute, V. J. (2008, March). Focus on formative feedback. *Review of Educational Research*, *78*(1), 153–189. Retrieved 2013-11-13, from http://www.jstor.org/stable/40071124 (ArticleType: research-article / Full publication date: Mar., 2008 / Copyright © 2008 American Educational Research Association)

Shute, V. J., & Armstrong Laboratory (U.S.). (1994). *Intelligent tutoring systems past, present,and future* (No. 1994-0005). Brooks Air Force Base, Tex: Armstrong Laboratory, Air Force Materiel Command.

Siemens, G., & Baker, R. S. J. d. (2012). Learning analytics and educational data mining: towards communication and collaboration. In *Proceedings of the 2nd international conference on learning analytics and knowledge* (p. 252–254). New York, NY, USA: ACM. Retrieved 2013-09-13, from `http://doi.acm.org.ezproxy.lib.utexas.edu/10.1145/2330601.2330661` doi: 10.1145/2330601.2330661

Smart machines in education: the coming revolution in educational technology. (2001). Menlo Park, Calif. : Cambridge, Mass: AAAI Press ; MIT Press.

Stevens, A. L., & Collins, A. (1977). The goal structure of a socratic tutor. In *Proceedings of the 1977 annual conference* (p. 256–263). New York, NY, USA: ACM. Retrieved 2013-11-13, from `http://doi.acm.org/10.1145/800179.810212` doi: 10.1145/800179.810212

Sweller, J. (2006, April). The worked example effect and human cognition. *Learning and Instruction*, *16*(2), 165–169. (WOS:000238018500007) doi: 10.1016/j.learninstruc.2006.02.005

Trivedi, S., Pardos, Z. A., & Heffernan, N. T. (2011). Clustering students to generate an ensemble to improve standard test score predictions. In G. Biswas, S. Bull, J. Kay, & A. Mitrovic (Eds.), *Artificial intelligence in education* (Vol. 6738, pp. 377–384). Berlin: Springer-Verlag Berlin. (WOS:000311898000049)

UMass amherst: Department of computer science. (n.d.). Retrieved 2013-09-24, from `https://web.cs.umass.edu/csinfo/announce/woolf.html`

VanLEHN, K. (2011). The relative effectiveness of human tutoring, intelligent tutoring systems, and other tutoring systems. *Educational Psychologist*, *46*(4), 197–221. Retrieved 2013-09-16, from `http://www.tandfonline.com/doi/abs/10.1080/00461520.2011.611369` doi: 10.1080/00461520.2011.611369

VanLehn, K., Siler, S., Murray, C., Yamauchi, T., & Baggett, W. B. (2003). Why do only some events cause learning during human tutoring? *Cognition and Instruction*, *21*(3), 209–249. (WOS:000184678400001) doi: 10.1207/S1532690XCI2103$_0$1

Viccari, R. M., Jaques, P. A., & Verdin, R. (Eds.). (2008). *Agent-based tutoring systems by cognitive and affective modeling:*. IGI Global. Retrieved 2013-09-13, from `http://www.igi-global.com.ezproxy.lib.utexas.edu/chapter/using-bayesian-networks-student-modeling/5043`

Vygotsky, L. S. (1978). *Mind in society: The development of higher psychological processes.* Harvard University Press.

Waalkens, M., Aleven, V., & Taatgen, N. (2013, January). Does supporting multiple student strategies lead to greater learning and motivation? investigating a source of complexity in the architecture of intelligent tutoring systems. *Computers & Education*, *60*(1), 159–171. (WOS:000312231900015) doi: 10.1016/j.compedu.2012.07.016

Wijekumar, K. K., Meyer, B. J., & Lei, P. (2013, October). High-fidelity implementation of web-based intelligent tutoring system improves fourth and fifth graders content area reading comprehension. *Computers & Education*, *68*, 366–379. Retrieved 2013-09-12, from `http://www.sciencedirect.com/science/article/pii/S0360131513001498` doi: 10.1016/j.compedu.2013.05.021

Wijekumar, K. K., Meyer, B. J. F., & Lei, P. (2012, December). Large-scale randomized controlled trial with 4th graders using intelligent tutoring of the structure strategy to improve nonfiction reading comprehension. *Etr&d-Educational Technology Research and Development*, *60*(6), 987–1013. (WOS:000311779000002) doi: 10.1007/s11423-012-9263-4

Woolf, B. P. (2009). *Building intelligent interactive tutors: student-centered strategies for revolutionizing e-learning.* Amsterdam ; Boston: Morgan Kaufmann Publishers/Elsevier.

Xu, F., & Tenenbaum, J. B. (2007, April). Word learning as bayesian inference. *Psychological Review*, *114*(2), 245–272. Retrieved 2013-09-12, from `http://ezproxy.lib.utexas.edu/login?url=http://search.ebscohost.com/login.aspx?direct=truedb=pdhAN=2007-05396-002site=` doi: 10.1037/0033-295X.114.2.245