

Project Summary

Exploring the Role of Emotion in Propelling the SMET Learning Process

Rosalind Picard, Barry Kort, Rob Reilly
Media Laboratory, M.I.T.
{picard, bkort, reilly}@media.mit.edu

The current educational system is constructed so as to emphasize the evolution of rule-based reasoners (students who understand the ‘what’ of a task/problem). However scientists, mathematicians, engineers, and technologists are not rule-based reasoners. They are model-based reasoners/thinkers (e.g., they do recursion, diagnostic reasoning, cognitive assessment, cognitive appraisal). To provide for the development of SMET (science, math, engineering, technology) learners, teachers and mentors who are reflective of the educational system’s pedagogy must not model directive teaching, which develops rule-based thinkers, but they should model how to learn. The educational pedagogy of a teacher and of the curriculum must include an understanding of learning intelligence—the ability to understand and model how-to-learn, how to solve a unique problem even when confronted with discouraging set-backs during its solution; in other words, they must be able to do model-based reasoning. Based upon this hypothesis, we propose to explore and evolve current educational pedagogy such that SMET learning will move toward a model-based reasoning pedagogy.

There has been a good deal of researcher conducted in this field that is very important. Their focus, however, is largely on how mood influences the content of what is learned or retrieved, and very much in a rule-based learning context. Our focus, in contrast, is on constructing a theory of affect in SMET learning, together with building practical tools that can specifically aid the student learn how to learn. We hypothesize that computers can begin to measure affect-related expression and behavior and can eventually become adept at adjusting the presentation by varying the pace, complexity, subtlety and difficulty

We are proffering a novel model by which to conceptualize the impact of emotions upon learning. We believe that there is an interplay of emotions and learning, but this interaction is far more complex than previous theories have articulated. Our model goes beyond previous research studies not just in the emotions addressed, but also in an attempt to formalize an analytical model that describes the dynamics of emotional states during model-based learning experiences, and to do so in a language that the SMET learner can come to understand and utilize.

We propose to discover, describe, and evolve the cognitive and affective learning processes required for SMET learning. We then propose to incorporate these research-based findings into a testbed simulation—the *Learning Companion* (a software-based interactive application) that will recognize the affective and cognitive state of the learner and respond in an appropriate manner (e.g., can adjust the pace, difficulty, complexity). Ultimately the *Learning Companion* will become an educational tool utilized in classrooms. We expect our results to be applicable to computer-based artifacts (e.g., our *learning companion*, companion-like software applets built into curricular software), and to impact the pedagogical approach of educators.

Project Description

Exploring the Role of Emotion in Propelling the SMET Learning Process

Rosalind Picard, Barry Kort, Rob Reilly
Media Laboratory, M.I.T.
{picard, bkort, reilly}@media.mit.edu

Why is there no word in English for the art of learning? Webster says that *pedagogy* means the art of teaching. What is missing is the parallel word for learning. In schools of education, courses on the art of teaching are simply listed as “methods.” Everyone understands that the methods of importance in education are those of teaching—these courses supply what is thought to be needed to become a skilled teacher. But what about methods of learning?

- Seymour Papert, *The Children’s Machine*

I. Introduction

Educators have traditionally emphasized conveying information and facts; rarely have they modeled the learning process. When teachers present material to the class, it is usually in a polished form that omits the natural steps of making mistakes (and feeling confused), recovering from them (overcoming frustration), deconstructing what went wrong (not becoming dispirited), and starting over again (with hope and maybe enthusiasm). Those of us who work in science, math, engineering, and technology (SMET) as professions know that learning naturally involves failure and a host of associated affective responses. Yet, educators of SMET learners have rarely illuminated these natural concomitants of the learning experience. The unfortunate result is that when students see that they are not getting the facts right (on quizzes, exams, etc.), then they tend to believe that they are either “not good at this,” “can’t do it,” or that they are simply “stupid” when it comes to these subjects. What we fail to teach them is that all these feelings associated with various levels of failure are normal parts of learning, and that they can be actually be helpful signals for *how* to learn better.

Scientists, mathematicians, engineers, and technologists tend not to be rule-based learners, who simply learn and apply facts, but rather model-based reasoners, who are capable of performing recursion, diagnostic reasoning, cognitive assessment, cognitive appraisal, and a host of other methods that require a real fortitude in learning ability. SMET learners routinely learn. Their knowledge is never sufficient: value can be gained from seeing even something they already understand in a new way. The educational system largely models directive teaching, which develops rule-based thinkers, and largely ignores the skills associated with learning intelligence—the ability to understand and model how-to-learn, how to solve a unique problem even when confronted with discouraging set-backs during its solution. The focus of this proposal is on evolving current educational pedagogy such that SMET learning will move toward a model-based reasoning pedagogy that includes the development of skills in learning intelligence, especially those skills that deal with the affective components of the learning experience. Toward this aim, we propose to develop, over several stages, a system that will be a *Learning Companion*.

The goal of building a computerized *Learning Companion* is to facilitate the child's own efforts at learning. Our initial aim will be to craft a companion that will help keep the child's exploration going, by occasionally prompting with questions or feedback, and by watching and responding to aspects of the affective state of the child—watching especially for signs of frustration and boredom that may precede quitting, for signs of curiosity or interest that tend to indicate active exploration, and for signs of enjoyment and mastery, which might indicate a successful learning experience. Although the *Learning Companion* may be teamed up with intelligent tutoring systems, it is not a tutor that knows the answers about the subject being learned. Instead, the *Learning Companion* will be a player on the side of the student—a collaborator of sorts—there to help him or her learn, and in so doing, learn how to learn better. It is a system that is sensitive to the learning trajectory of students, and that, we hypothesize, will in turn increase the sensitivity of learners to their own learning trajectory. It will have succeeded if students, especially those who encounter frustration and routinely handle it by quitting, learn instead how to persevere, increasing their ability and desire to engage in self-propelled learning.

The *Learning Companion* will serve as a test bed for three concurrent areas of research we are working on related to affect and learning: 1. Understanding which emotions are most important in learning; 2. Evolving different methods for computers to use in recognizing affective states important to learning; and, 3. Building and evaluating different strategies of learning pedagogy related to student awareness of affective states. In short: what emotions are important, how can they be detected?, and how should the companion respond to them? Below, we present several key ideas related to these areas, highlighting how current emotion theories have ignored emotions specific to learning, reviewing how technological advances are increasing abilities of computers to infer users’ affective expressions, and describing an initial strategy we propose to test for how affect can influence

learning intelligence. We believe that each of these areas, while well-grounded in prior work, is also subject to a learning process; thus, the ideas below are in the spirit of starting with something substantial, while allowing room for variation as we learn which aspects of the companion are most successful as we build it and test it with kids.

II. Affective Computing: Emotions and Learning

The extent to which emotional upsets can interfere with mental life is no news to teachers. Students who are anxious, angry, or depressed don't learn; people who are caught in these states do not take in information efficiently or deal with it well.

Daniel Goleman, *Emotional Intelligence*

Expert teachers are very adept at recognizing and addressing the emotional state of learners and, based upon that observation, taking some action that positively impacts learning. But what do these expert teachers 'see' and how do they decide upon a course of action? How do they return the student to the 'zone of flow' when they have strayed?

Preliminary research by Lepper and Chabay [1988] indicates that "expert human tutors... devote at least as much time and attention to the achievement of affective and emotional goals in tutoring, as they do to the achievement of the sorts of cognitive and informational goal that dominate and characterize traditional computer-based tutors." We propose to further examine what expert teachers 'see' as well as examine what they do in response to this and integrate these findings into the *Learning Companion*.

Skilled humans can assess emotional signals with varying degrees of accuracy, and researchers are beginning to make progress giving computers similar abilities at recognizing affective expressions. Although computers perform as well as people only in highly restricted domains, we believe that accurately identifying a learner's emotional/cognitive state is a critical indicator of how to assist the learner in achieving an understanding of the efficiency and pleasure of the learning process. We also assume that computers will, much sooner than later, be more capable of recognizing human behaviors that lead to strong inferences about affective state. We review here some of the different means that computers can use to assess affective state. We propose to continue evolving these methods, in the context of a *Learning Companion*, to evaluate which methods are most comfortable and helpful for the students.

Methods of Inferring Affective State

Questionnaires: Matsubara and Nagamashi [1996] employed questionnaires in the beginning of interactions "to diagnose several factors influencing motivation, such as: Achievement Motive, Creativity, Sensation Seeking Scale, Extroversion-Introversion, Work Importance and Centrality, Least Preferred Co-worker and Locus of Control" [de Vincente and Pain, 1998]. Whitelock and Scanlon [1996] used post-test questionnaires to assess a number of affective factors such as "curiosity, interest, tiredness, boredom and expectation plus the challenge of the task" to assess the affective state. Klein *et al.* have developed dialogue boxes with radio buttons for querying users about emotion [Klein, *et al.* 1999], not only to identify the frustration level of a user, but also to tailor a response to it. Their study with 70 subjects showed that a particular "active listening" style of response led to a significant decrease in user frustration, based on a behavioral measure and comparison with two control groups. Thus, an interactive questionnaire might not only help assess emotion, it might also help the user better manage their emotions.

Pre-interaction questionnaires have been criticized for being static and thus not able to recognize changes in affective states during research interactions. Questionnaires are also subject to a common problem that plagues all methods of self-report: the social-emotional expectations and awareness of the subject can greatly influence what is reported. For example, a subject who thinks it is bad to feel angry in a classroom may never report that something angered them. On the other hand, questionnaires are an easily administered means for detecting affective states and several have been devised to detect motivation state change [Gardner, 1985]. We propose, similar to de Vincente and Pain [1998] to "use questionnaires for collecting information about enduring characteristics of the student that can help to adapt instruction, although other methods should be used to gather information about more transient characteristics."

Help-based interactions: del Soldato [1994] had success in gathering information about the subject's affective state via face-to-face dialogue, e.g., direct conversation with the student in regard to their affective state, during the treatment, and during the student's request for help/assistance. Although many students are willing to ask for help, the assumption that all students are able, and/or willing to ask for help, is a serious shortcoming of educational software. Studies of spoken assistance on demand (Olson et al 1986, Olson and Wise 1987, Conkie 1990, Olofsson 1993) have revealed a serious flaw in assuming that young readers are willing and able to ask for help when they need it. And there is the problem that the student does not know when they are in-trouble (children with reading difficulties often fail to realize when they misidentify a word). This is especially acute for children with weak metacognitive skills. The stigma of being "thought stupid" also prevents many kids from asking questions;

we think this feeling relates to a lack of comfort with the negative emotions such as confusion and frustration, and to a lack of understanding of the important and essential roles such feelings play during many a challenging learning episode.

Self-Report: Self-report methods include questionnaires and interviews that might be conducted briefly during a help session, as in the two cases above. (Although interviews can still be better for assessing the non-verbal aspects of the self-report.) Self-report tools can also include special buttons and sliders. del Soldato [1994] theorized about the use of special buttons and on-screen 'sliders' in the design of the MORE system. Keller's [1987a] ARCS model, which is rooted in a number of motivational theories and concepts, (see Keller, 1983) most notably expectancy-value theory (e.g. Vroom, 1964; Porter and Lawler, 1968), identifies four components that appear to affect motivation to learn: attention, relevance, confidence, satisfaction (hence, ARCS). Creatively combining the work of the aforementioned authors, an "interface could easily be implemented with mechanisms that would allow the student to report his/her subjective reading of these factors. For example, each of these factors could be represented by a slider, which could be manipulated by the student" [de Vincente and Pain, 1998]. We plan to evaluate the desirability of including such a mechanism(s) in the *Learning Companion*. Our evaluation will be accomplished by utilizing several of the instruments that have been developed for assessing the motivational quality of instructional situations. Such instruments include the Instructional Materials Motivation Survey [Keller, 1987], which asks students to rate 36 ARCS-related statements in relation to the instructional materials they have just used., or the Motivational Delivery Checklist [Keller and Keller, 1989], which is a 47-item ARCS-based instrument for evaluating the motivational characteristics of an instructor's classroom delivery.

While self-reports are quite easy to implement, they would also seem to be the least reliable given that the subjects may not be able to perform an accurate 'in flight' diagnosis of their state while focusing on the primary task, or the subject may simply not be able to accurately identify one affective state from another [Briggs, 1996]. Other researchers have concluded that "It is not clear [from the available research] whether or how the student's motivational state will be affected by having to report about" it [de Vincente and Pain, 1998]. In the general emotion theory literature, self-report methods are known to be rather unreliable compared to behavioral measures such as "how long did the subject choose to engage with the system after his or her commitment was over?" We propose to look jointly at both self-report data and at other behavioral variables, since the combination of measures can give stronger results. We also propose to conduct interviews with users of the learning companion to assess their feelings, before and after, about the effects of the system on their emotional awareness and how it may have helped or hindered the learning experience. We will also plan to interview teachers of these students for their feedback.

Expert System: A prototype expert system developed by Hioe and Campbell [1988] "assisted managers in finding which problems were affecting employee performance." The system reviewed various human motivation theories. Then based upon the expert's diagnostic processes, these responses were placed into four groups and then further processed. Although this was not a computer tutor, it is not difficult to envision a similar approach utilized in an intelligent tutoring system or in a learning companion. While this approach would not be as dynamic as other approaches (e.g., sentic modulation, below) it would neither be as static as a questionnaire.

Sentic Modulation: This area is in its infancy but it ultimately offers the most dynamic and objective approach for assessing changes in a person's affective state. Sentic modulation refers to the physical assessment of a person's emotional changes via sensors such as cameras, microphones, strain gauges applied to mouse buttons, special wearable devices, and other indicators that register subtle physical modulation produced by emotional states. The basic idea is that an emotional state modulates many physical changes in tandem: it may increase or decrease tension in muscle groups, provoke fewer or greater specific facial actions, slow down or speed up eye and finger movement, shift directionality of movement (say, toward the item of liking or away from the item of disliking), increase or lower electrodermal responses, and so forth. The job of the computer is to assess a constellation of such patterns and relate them to the user's affective state.

Most prior work on emotional expression recognition from speech, image, and video has focused on deliberately expressed emotions, not on those that occur in natural situations such as classroom learning. The results make it hard to predict rates we can expect when relating emotions to learning. In general, people can recognize emotion in neutral-content speech with about 60% accuracy, choosing from among about six different affective states [Scherer, *et al.*, 1999]. Computer algorithms match this accuracy under more restrictive assumptions, such as when the sentence content is already known. However, automated speech recognition that works at about 90% accuracy on neutrally spoken speech tends to drop to 50-60% accuracy on emotional speech [Hansen 1999]. Improved handling of emotion in speech is important for improving recognition of *what* was said, as well as *how* it was said. Facial expression recognition is easier for people, and the rates computers obtain are higher: from 80-98% accuracy when recognizing 5-7 classes of emotional expression on groups of 8-32 people [Yacob and Davis, 1996; Essa 1997]. Other research has focused not so much on recognizing a few categories of "emotional expressions" but on recognizing specific facial actions—the fundamental muscle movements that comprise Paul Ekman's Facial Action Coding System, which can be combined to describe all facial expressions [Ekman, 1977]. Recognizers have already been built for a handful of the facial actions [Essa, 1995; Cohn *et al.*, 1999; Bartlett *et al.*, 1999; Donato *et al.*, 1999] and the automated recognizers have been shown to perform comparably to humans trained in recognizing facial actions [Cohn *et al.*, 1999]. These facial actions are

essentially *facial phonemes*, which can be assembled to form facial expressions. Other recent studies indicate that combining multiple modalities, namely audio and video, for emotion recognition can give improved results [DeSilva *et al.*, 1997; Huang *et al.*, 1998; Chen *et al.*, 1998].

Although the progress in facial, vocal, and combined facial/vocal expression recognition is promising, the results above are on pre-segmented data of a small set of sometimes exaggerated expressions, or on a small subset of hand-marked singly-occurring facial actions. The state of the art in affect recognition is similar to that of speech recognition several decades ago when the computer could classify the carefully articulated digits, “0, 1, 2, ..., 9,” spoken with pauses in between, but could not accurately detect these digits in the many ways they are spoken in larger continuous conversations. Thus we cannot expect the computer to perform perfectly at recognition, and our methods will have to take into account uncertainty factors.

Picard and her students have been actively engaged in building systems that sense and try to interpret patterns of sentic modulation and their relation to underlying emotional states [Picard, 1997]. Her group has focused on computer recognition of truly felt emotions, as opposed to merely those that have been expressed by actors or by subjects posed in front of a camera or microphone. Recent systems developed in the MIT Media Laboratory include “expression glasses,” which discriminate upward facial expressions such as those of interest and openness from downward expressions such as those of confusion or dissatisfaction [Scheirer *et al.* 1999], and a physiological monitoring system that senses four signals from the surface of the skin and relates these to eight emotional states (subject-dependent) with over 81% classification accuracy [Vyzas and Picard, 1999]. Recent findings reported by Healey [2000], measuring physiological changes due to increased workload and stress in drivers on Boston roads, showed recognition rates of 96% based on three types of situational stress, and of 89% based on self-reports of stress.

Picard’s group has also developed surface sensors of mouse-clicking information and video-based analysis tools for other forms of human behavior. We propose to develop additional computer-vision and computer-audition means of looking at and listening to learners in an unobtrusive way. We also propose to integrate these artifacts into the Learning Companion, so that it can begin to tailor its responses not just to cues like number of errors made by the learner, but also to affective cues, such as “did the learner make these errors with a look of curiosity and interest?” or “did the learner make them with mannerisms that look increasingly frustrated and angry?” The ability to recognize such affective cues by computer is an extremely challenging area of basic research, but one which we believe will benefit from and will add benefit to the development of a Learning Companion.

III. Related research

Our goal is to transform how children learn, what they learn, who they learn from.

-Mitchel Resnick, *A Media Lab for Kids*

One of the first suggestions for endowing computer tutors with a degree of empathy or affect was made by Lepper and Chabay [1988]. “They argued that motivation components are as important as cognitive components in tutoring strategies, and that important benefit would arise from considering techniques to create computer tutors that have an ability to empathise” [de Vincente and Pain, 1998]. Others such as Issroff and del Soldato [1996] and Lepper and Chabay [1988] have suggested that Intelligent Tutoring Systems (ITS) should include a mechanism for motivating the learner, detecting a learner’s emotional/motivational state and responding to that state.

Alice Isen [1999] at Cornell, the editor of *Motivation and Emotion* has researched the relationship between mood, creativity, and problem solving. Her work in marketing has focused on how putting customers, clients, or medical patients in a good mood causes changes in their subsequent cognitive and problem solving abilities, generally improving these abilities significantly as compared to controls. Niedenthal’s [1999] work dealing with how emotions bias perception and categorization is also of interest to us, highlighting interactions between what is perceived and the affective state of the perceiver. There have also been many research efforts to associate emotional and cognitive states and their influence on learning; a survey by Mayer [1986] highlights issues involved in reproducing the classical experimental results of Bower and Cohen, showing that mood repeatedly influences judgment and, under certain controls, the nature of the learning experience.

Csikszentmihalyi [1990] and Vail [1994] have delved into the idea of emotions and learning. Vail has compiled a great deal of practical advice that is consistent with the theoretical model that we propose below. Csikszentmihalyi [1990] has come the closest to building a theoretical model that incorporates affect.

We believe that the researchers cited above and the findings they report are important. Their focus, however, is largely on how mood influences the content of what is learned or retrieved in a rule-based learning context. Other researchers have explored how various human personality types respond to computers (i.e., Reeves 1996), which is also relevant in crafting a companion that

will facilitate an individual's learning. Still, "the very idea that computer systems are able to address and respond to the user's emotional state represents an important departure in human-computer interaction (HCI) practice" [ITS Working Group, 1995c]. Beyond some work at the periphery of the field, current theory in HCI reflects a lack of direct consideration for the user's emotional state. When reading current HCI literature, for example, "it seems doubly strange that this important aspect of the user's experience continues to be largely unexplored" [Klein, 1996b].

Applying Related Research—Exploration

In the exploratory phase of this research, we conducted a pilot experiment, the outcome of which influenced our current plans. This research involved setting up an unobtrusive camera and getting permission from parents to videotape and observe 30—2nd grade and 4th grade (8-9 and 10-11 year old) students in an actual learning situation attempting to solve the various puzzles in *The Incredible Machine*. The students self-selected themselves into groups of one, two or three and then ran the software for 20 minutes. We observed the hours of video that were collected, looking for various styles of interaction and the correlation of these styles with progress toward the puzzle-solving goals.

The most salient feature from all the interactions was that students who worked alone tended to become hopelessly lost more often than the two-student groups, and students who worked in groups of three engaged in off-goal-behaviors (largely socializing, instead of trying to work on the puzzles) to a great degree. During subsequent trials where the students selected other groupings, it was interesting to note that the students who were previously in groups of three attended to the task much better when they were in groups of two, and those who were previously in groups of two became hopelessly lost more readily when working alone. It appeared that the students when working in groups of two provided mutual support to work through the more difficult passage—a collaboration effect. Even though the companion student in these groups of two was typically not intellectually superior to the other, the reinforcement provided by the companion facilitated successful problem-solving or, at least significantly forestalled 'quitting.'

One of the key parts of the learning process, which is often overlooked, is the sheer fact of keeping learning going after the student feels like quitting. The presence of a learning companion (in this case a human companion—a collaborator—a peer) was a significant factor in how long the students persevered in the technical and creative problem solving involved in *The Incredible Machine*.

This positive effect of a peer is not a new finding; various studies have demonstrated that 'collaboration' with a peer is beneficial to enhancing and/or increasing learning [Azmitia, 1988; Doise *et. al.*, 1975, 1976; Ellis, *et. al.*, 1993]. Other studies have examined the types of supportive conversation that accompany computer supported collaborative learning activities [Whitelock *et. al.*, 1993, Mercer 1994, Wegerif 1996]. These and still other studies with children have also found that peer presence facilitates problem solving [Joiner *et. al.*, 1991] and that gender too has a mediating effect [Loveridge *et. al.*, 1993]. The benefits of peer-presence are not universal and may vary across tasks and with individual students [e. g., Tudge, 1989]; nonetheless, the students seem to learn more effectively when they collaborate with their peers. This has been shown to be especially effective when the task is conceptual and complex [Gabbert, *et. al.*, 1986], as one typically finds in SMET learning. Collaboration with a peer also seems to have other beneficial effects such as improving social relations, or increasing students' motivation [Sharan, 1980]. Thus, collaboration can be viewed as an effective instructional medium. However, there has been little principled investigation into the role of the non-verbal interactions which accompany and support such cognitive skills as planning and problem solving within a computer supported learning situation where a peer is present.

While researchers have proposed that several factors such as cognitive conflicts [Doise *et. al.*, 1975, 1976], partner expertise [Azmitia, 1988], or increased amount of verbalization [Teasley, 1992] are responsible for improving learning in collaboration, these factors do not provide an explanation of how collaboration actually works.

By building a learning collaborator—a *Learning Companion*, which the students would perceive as being a collaborator—being 'on their side' in trying to help them in a learning task—we therefore build a test environment for more principled exploration of the verbal and non-verbal communication that happens between a learner and a collaborator. Although the computer companion won't be as intelligent as a human companion, we observed that the human companions present for kids playing the Incredible Machine rarely engaged in significant discourse, but mostly engaged in pointing, posing questions, and re-directing attention when the learner seemed to slow down or appear stuck or frustrated. Thus, we are inspired to see what a computerized companion could do if it could attend not only to the traditional cognitive factors involved in learning (making of errors, hesitations in acting, etc.), but also to attend to affective factors (signs of frustration vs. interest, etc.). For example, the companion would not only notice if the student is making errors, but also if the student is doing so with curiosity and interest, or with increasing frustration. Different responses of the companion could be crafted and tested for these different cases – offering more encouragement or perhaps not intervening at all, adjusting subtlety of feedback and other strategies, depending on the

sensed conditions. The space of possibilities here is quite complex, and we outline below some more particular affective states and scenarios that we think are significant to explore based on our combined experience of over 80 years of SMET learning.

IV. The Guiding Theoretical Frameworks—Developing an Advanced Technology

In order to accomplish our goal we must redefine and in some cases reengineer various aspects of educational pedagogy. To this end it is necessary for us to briefly present our perspective of what is happening in education and what we believe must occur. Some of these beliefs will be theorized, perhaps beyond a practical level but not beyond a level needed for understanding them. We need to explore the underpinnings of various educational theories and evolve or revise them. For example, we propose a model that describes the range of various emotional states during learning (see Figures 1, 2a, 2b). The model is inspired by theory often used to describe complex interactions in engineering systems, and as such is not intended to explain how learning works, but rather is intended to give us a framework for thinking about and posing questions about the role of emotions in learning. Like with any metaphor, the model has limits to its application. In this case, the model is not intended to fully describe all aspects of the complex interaction between emotions and learning, but rather only to serve as a beginning for describing some of the key phenomena that we think are all too often overlooked in learning pedagogy. Our model goes beyond previous research studies not just in the emotions addressed, but also in an attempt to formalize an analytical model that describes the dynamics of emotional states during model-based learning experiences, and to do so in a language that the SMET learner can come to understand and utilize.

An Affective Learning System Model

Pedagogy, the art of teaching, under various names, has been adopted by the academic world as a respectable and an important field. The art of learning is an academic orphan. One should not be misled by the fact that libraries of academic departments of psychology often have a section marked “learning theory.” The older books under this heading deal with the activity that is sometimes caricatured by the image of a white-coated scientist watching a rat run through a maze...newer volumes are more likely to be based upon the theories of performance of computer programs than on the behavior of animals... but... they are not about the art of learning... they do not offer advice to the rat (or to the computer) about how to learn.

- Seymour Papert, *The Children’s Machine*

Before describing the model’s dynamics, we should say something about the space of emotions it names. Previous emotion theories have proposed that there are from two to twenty basic or prototype emotions (see for example, Plutchik, 1980; Leidelmeijer, 1991). The four most common emotions appearing on the many theorists’ lists are fear, anger, sadness, and joy. Plutchik [1980] distinguished among eight basic emotions: fear, anger, sorrow, joy, disgust, acceptance, anticipation, and surprise. Ekman [1992] has focused on a set of from six to eight basic emotions that have associated facial expressions. However, none of the existing frameworks seem to address emotions commonly seen in SMET learning experiences, some of which we have noted in Figure 1. Whether all of these are important, and whether the axes shown in Figure 1 are the “right” ones remains to be evaluated, and it will no doubt take many investigations before a “basic emotion set for learning” can be established. Such a set may be culturally different and will likely vary with developmental age as well. For example, it has been argued that infants come into this world only expressing interest, distress, and pleasure [Lewis, 1993] and that these three states provide sufficiently rich initial cues to the caregiver that she or he can scaffold the learning experience appropriately in response. We believe that skilled observant human tutors and mentors (teachers) react to assist students based on a few ‘least common denominators’ of affect as opposed to a large number of complex factors; thus, we expect that the space of emotions presented here might be simplified and refined further as we tease out which states are most important for shaping the companion’s responses. Nonetheless, we know that the labels we attach to human emotions are complex and can contain mixtures of the words here, as well as many words not shown here. The challenge, at least initially, is to see how much a companion can do initially with a very small space of possibilities, since the smaller the set, the more likely we are to have greater classification success by the computer.

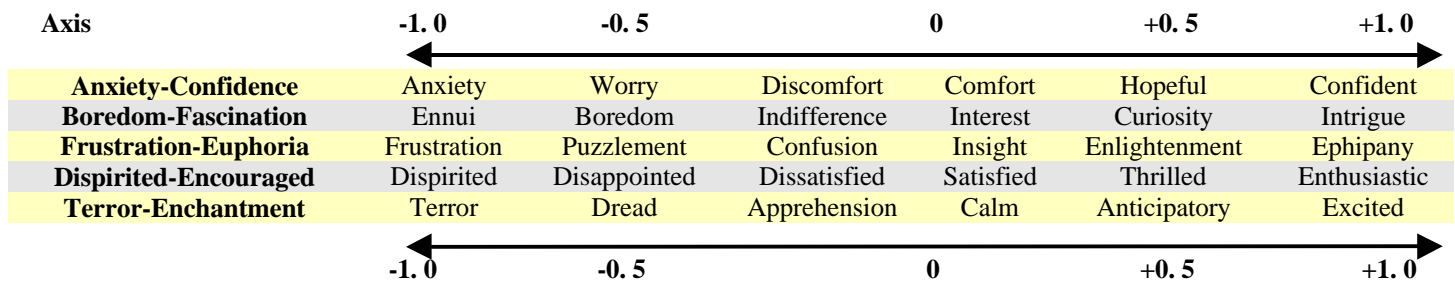


Figure 1 – Emotion sets possibly relevant to learning (in contrast to traditional emotion theories)

Figure 2 attempts to interweave the emotion axes above (Figure 1) with the cognitive dynamics of the learning process. The horizontal axis is an Emotion Axis. It could be one of the specific axes from Figure 1, or it could symbolize the n-vector of all relevant emotion axes (thus allowing multi-dimensional combinations of emotions). The positive valence (more pleasurable) emotions are on the right; the negative valence (more unpleasant) emotions are on the left. The vertical axis is what we call the Learning Axis, and symbolizes the construction of knowledge upward, and the discarding of misconceptions downward. (Note that we do **not** see learning as being simply a process of constructing/deconstructing or adding/subtracting information; this terminology is merely a projection of one aspect of how people can think about learning. Other aspects could be similarly included along the Learning Axis.)

The student ideally begins in quadrant I or II: they might be curious and fascinated about a new topic of interest (quadrant I) or they might be puzzled and motivated to reduce confusion (quadrant II). In either case, they are in the top half of the space, if their focus is on constructing or testing knowledge. Movement happens in this space as learning proceeds. For example, when solving a puzzle in *The Incredible Machine*, a student gets an idea how to implement a solution and then builds its simulation. When she runs the simulation and it fails, she sees that her idea has some part that doesn't work – that needs to be deconstructed. At this point it is not uncommon for the student to move down into the lower half of the diagram (quadrant III) where emotions may be negative and the cognitive focus changes to eliminating some misconception. As she consolidates her knowledge—what works and what doesn't—with awareness of a sense of making progress, she may move to quadrant IV. Getting a fresh idea propels the student back into the upper half of the space, most likely quadrant I. Thus, a typical learning experience involves a range of emotions, moving the student around the space as they learn.

If one visualizes a version of Figures 2a and 2b for each axis in Figure 1, then at any given instant, the student might be in multiple quadrants with respect to different axes. They might be in quadrant II with respect to feeling frustrated; and simultaneously in quadrant I with respect to interest level. It is important to recognize that a range of emotions occurs naturally in a real learning process, and it is not simply the case that the positive emotions are the good ones. We do not foresee trying to keep the student in quadrant I, but rather to help them see that the cyclic nature is natural in SMET learning, and that when they land in the negative half, it is only part of the cycle. Our aim is to help them to keep orbiting the loop, teaching them how to propel themselves especially after a setback.

A third axis (not shown), can be visualized extending out of the plane of the page—the Knowledge Axis. If one visualizes the above dynamics of moving from quadrant I to II to III to IV as an orbit, then when this third dimension is added, one obtains the “excelsior spiral that climbs the tree of knowledge.” In the phase plane plot, time is parametric as the orbit is traversed in a counterclockwise direction. In quadrant I, anticipation and expectation are high, as the learner builds ideas and concepts and tries them out. Emotional mood decays over time, either from boredom or from disappointment. In quadrant II, the rate of construction of working knowledge diminishes, and negative emotions emerge as progress flags. In quadrant III, the learner discards misconceptions and ideas that didn't pan out, as the negative affect runs its course. In quadrant IV, the learner recovers hopefulness and positive attitude as the knowledge set is now cleared of unworkable and unproductive concepts, and the cycle begins anew. In building a complete and correct mental model associated with a learning opportunity, the learner may experience multiple cycles around the phase plane until completion of the learning exercise. Each orbit represents the time evolution of the learning cycle. Note that the orbit doesn't close on itself, but gradually moves up the knowledge axis.

It is also possible to diagram Knowledge, Learning, or Emotions in a conventional Cartesian graph against time (again, a vast simplification of these complex concepts, but one that is helpful for thinking about their relationships.) We suggest an analogy between such a diagram and plotting Distance, Velocity, or Acceleration against Time in a traditional application of the Calculus (see Figures 3 and 4). The *roller coaster* plots would reflect the ebb and flow of emotions as learning proceeds with its cyclical gains and losses. Of the various ways to diagram the interplay of emotions and learning, the greatest insight comes from the 3-dimensional helix when Emotions and Learning are plotted in the phase plane, and cumulative Knowledge is plotted in the Z-axis, perpendicular to the Emotion-Learning phase plane. These models provide a framework for observing, analyzing, and

interpreting the emotion cycles that accompany learning, for diagnosing phases where the learner is stuck, frustrated, bored, or otherwise dispirited, and for crafting intelligent interventions to keep the learner moving optimally along the learning helix.

A computer *Learning Companion* could potentially use models such as these (which are standard engineering models) to assess whether or not learning is proceeding at a healthy rate. The models could help guide it in exploring strategies for making decisions about when best to intervene with a hint, word of encouragement, or observation (typically in quadrants III and IV.) Thus, we see the *Learning Companion* helping scaffold the learning experience by trying to keep the learner moving through this space, e.g., not avoiding quadrant III, but helping them to keep moving through it instead of getting stuck there. The models may also be useful to learners in aiding in their own metacognition about their learning experience, especially helping them identify and work with naturally-occurring negative emotions in a productive and cognitively satisfying way.

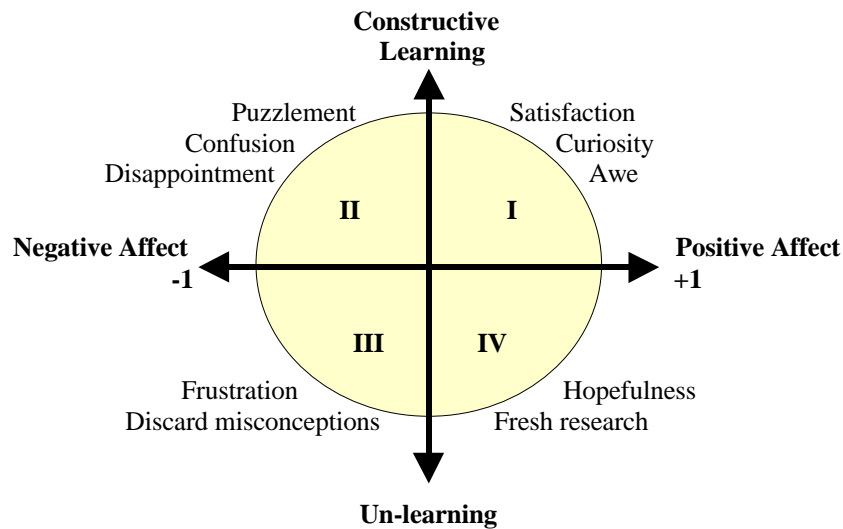


Figure 2a – Proposed model relating phases of learning to emotions in Figure 1

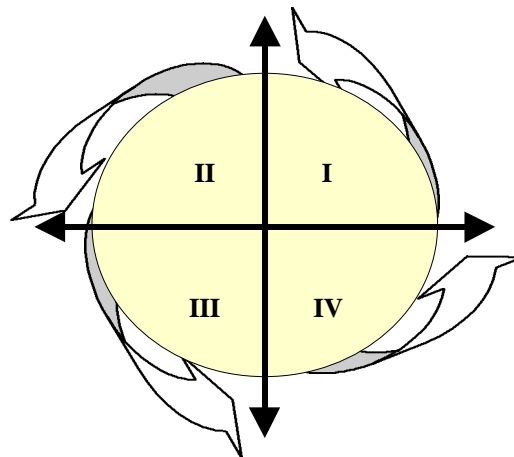


Figure 2b – Circular and Helical Flow of Emotion Through the Learning Journey

Current technological expertise cannot reliably recognize all the categories of emotions discussed above (Figure 1); however, we expect that it may not be essential to identify the exact emotional state continuously throughout the learning experience. What may be most important is that we are able to identify a change in emotion, which causes a learner to go from an *on-goal* state to *off-goal* state. We hypothesize that being able to identify a change at the lowest signal level is a critical indicator for intelligent intervention by our *Learning Companion*. Perhaps we only need to know that a learner is in the cognitive assessment or cognitive appraisal state in order to assist the learner. These are research issues that we propose to explore during this work.

Although the model as described uses more complex emotional descriptions than one can currently sense reliably with technology, it is still possible to use the model in these two ways: 1. with a reduced, less complex space of emotions, as just mentioned; and 2. with “a man behind the curtain” who conducts more complex emotion recognition, but lets the model drive the responses. Thus, we foresee at least these two paths to explore the utility of the model, even though at this point it is not possible to fully implement all of the model’s features. We propose to systematically explore the model via these two simpler paths, initially, and via more complex paths as the technology allows and as the model’s worth is proven (if it is proven) or as it is adapted (if an alternate formulation is needed to better fit the behavior of learners.)

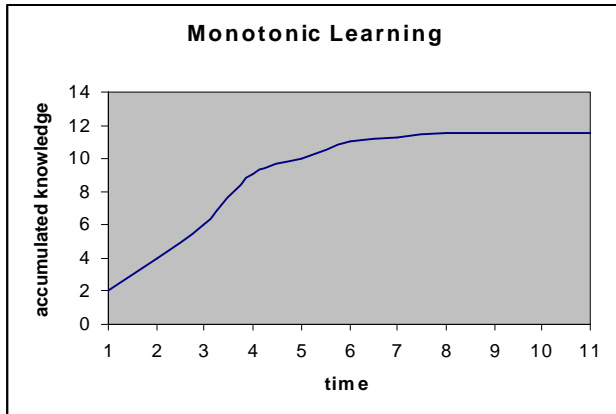


Figure 3 – Hypothetical Monotonic Learning Curve

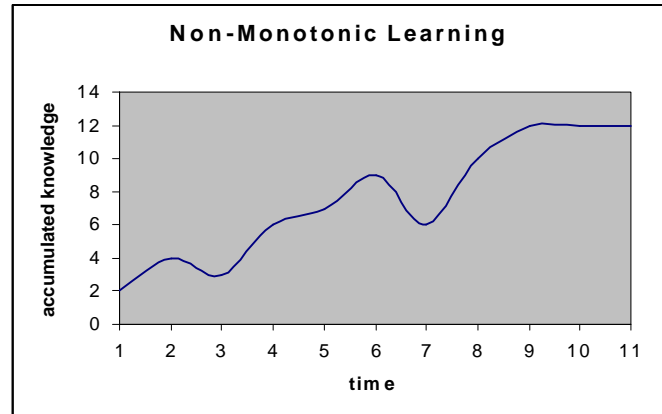


Figure 4 – Hypothetical Non-Monotonic Learning Curve

Another Ingredient: Scaffolding—A Supporting-Fading Mechanism for the “Learning Companion”

We have only begun to explore what are the appropriate scaffolds for promoting learning. We have also much to learn on how computational and communication technologies can support teacher collaboration and professional development. Although we are encouraged by our successes to date, we recognize that we are still in the early stages of a substantial effort.

- Eliot Soloway, *Scaffolding Technology Tools to Promote Teaching and Learning in Science*

In “designing software for education, [there is a conscious effort to] design [it] for learners... [and the] Highly Interactive Computing [(Hi-C)] group at the University of Michigan has formulated a rationale for learner-centered design (LCD) [Soloway, 1999]. “However, user-centered design guidelines are not sufficient to address certain unique needs of learners, such as intellectual growth, diversity of learning styles, and motivational needs” [Soloway 1999]. For example, “learners should have software available to them that represents information in a familiar way, but that also helps introduce them to more professional or symbolic representations” [Soloway, 1999]. The central claim of the Hi-C group is that the software can incorporate learning supports—*scaffolding*—to address the learner’s needs. Scaffolding is important as it enables the learner to achieve goals and/or accomplish processes that would not normally be possible or would be normally out of reach [Vygotsky, 1978; Wood, *et. al.*, 1975].

Soloway *et. al.* [1996] and Jackson *et. al.* [1999] have suggested the concept of scaffolding (fadable supports) designed into educational software to provide support “so that the learner can engage in activities that would otherwise be beyond their abilities” [Jackson *et. al.*, 1999]. A critical component of scaffolding is that it be capable of fading; as the learner’s understanding and abilities improve, the computer, much like a human tutor, needs to back off and give the learner more autonomy, fewer hints, etc. In the field of educational software research, scaffolding is a new concept that is still being defined [Collins, 1996; Guzdial, 1995; Soloway *et al.*, 1994; Soloway *et al.*, 1996]. We believe that this approach has great merit and our proposal will address the open question that it may be difficult for a learner to properly make fading decisions.

We surmise that a support mechanism for any learning experience needs to be one that can come-and-go as needed—not just fade. At different times a learner will require various levels of assistance (e.g., hints, outright answers, review of previous learning). At various times a learner will also require no assistance, and the assistance mechanism will then lay dormant. Thus the *Learning Companion* will need a mechanism to provide scaffolding—a method to provide a supporting-fading mechanism for each element of the learning environment. However, the focus of our investigation will be on which affective cues the learning companion can use to best tailor the scaffolding, and thereby explore how to maximize its effectiveness.

Another Ingredient: Privacy and Confidentiality

Privacy/confidentiality is the issue of the 21st century just as ‘Civil Rights’ was the issue of the ‘60s and ‘70s. Privacy has been a concern for a majority of Americans since the 1970. The ability to invade a person’s privacy to gather mounds of personal information has been growing due to increasing technological sensing capabilities. This issue will become even more insidious as our educational programs are becoming more capable of invading our privacy by providing intimate details of our every keystroke. Privacy/confidentiality and the unintended power acquired by creating the code for the *Learning Companion*, while not exclusively an affective issue, is none the less an important factor in creating a technological artifact [Lessig, 1999; Reilly, 1999]. The Affective Computing group at MIT has a policy on respecting user privacy (see www.media.mit.edu/affect) and we will be open with all subjects as to what the capabilities of the system are, giving users maximal control over the system and complete control over any data gathered by the system. All of our work with subjects is also conditional upon approval by the MIT Committee on the Use of Humans as Experimental Subjects, from whom we already gaining approval for the pilot study described above, and for several other experiments that sense and respond to user emotions in natural situations.

V. Evolving an Affective Learning System—The “Learning Companion”

Leaps in Technology

It may seem that the complexity of our emotional lives—our emotions—may be such that they cannot be captured or transferred to a computer and used for learning. But such questions have been commonplace when new technology is about to be born. When new technology has proposes to supplant an existing one there are always questions as to whether it can be done. For example, when steam engine powered ships were introduced with the intent of replacing sail-powered craft, they were not models of success—they routinely exploded. But eventually, steam ship technology was refined; and ultimately the steam ship too was replaced. With this said, let us proceed to outline this study.

Alongside and Beyond Intelligent Tutoring Systems

Today’s ITSs are hand-crafted, monolithic, standalone applications. They are time-consuming and costly to design, implement and deploy. Each development team must redevelop all of the component functionalities [the modules] needed. Because these components/[modules]are so hard are costly to build, few tutors of realistic depth and breadth ever get built, and even fewer ever get tested on real students. This is a Bad Thing.

- ITS Working Group, *Working Group: Exploring Industry Standard Architectures*

When first introduced more than 10 years ago Intelligent Tutoring Systems (ITSs), which were the next generation of Computer Assisted Instruction systems, “were avowed as the future of education and training” [Jerinic and Devedzic, 2000]. Despite some initial successes (Anderson , 1990; Bonar, 1988; Russell *et. al.*, 1988; Sleeman, 1987; Woolf, 1987), “ITSs have not yet seen general acceptance” [Jericin and Devedzic, 2000]. And today the ITS community “is still talking about the promise of this technology while searching for the leverage that will encourage its widespread adoption and classroom use” [Jericin and Devedzic, 2000]. The difficulties seem to be rooted in the definition and design complexities and pedagogical changes involved in ITS applications. We believe that ITSs, while producing visually appealing and interactive screens for presenting that first layer of information, only begin to scratch the surface of what can be done with learning pedagogy.

To review briefly, ITSs are computer-based instructional systems that are composed of separate knowledge bases, or data bases, for the various instructional content, for teaching strategies, individual user data, error correction, etc. The various modules attempt to use inferences about a student's mastery of topics to dynamically adapt instruction or to provide substantive information/knowledge. Woolf [1992] observes that typical ITSs consists of:

- a **Domain Knowledge Module**, which contains the information that the tutor is teaching,
- an **Expert Module**, is more than a mere representation of the data, it is a model of how an expert human teacher would present the Domain Knowledge,
- a **Student Module**, which maintains information that is specific to each user (e.g., how far they have progressed),
- a **Tutor/Pedagogical Module**, which is responsible for deciding how and when the domain knowledge is presented; this module emulates the pedagogical approach of an expert teacher (e.g., when to present a new topic, which topic to present, when a review is needed),
- a **Diagnostic/Misconception Module**, which contains the rules used to identify misperceptions and misunderstandings on the part of the user,

- a **Communication Module**, which is the user interface (e.g., keyboard, mouse, sentic device, screen display/layout). This module answers the question: “How best to present the material to the learner?” and,
- Jerinic and Devedzic [1997] have added an **Explanation Module**, which is “define[d as] the contents of explanations and justifications of the ITS’s learning process, as well as the way they are generated” (e.g., canned text, templates more fully explaining concept *x*, presentation models to explain concept *x* in concert with other modules).

Taking into account that prototypes are built incrementally through successive enhancements and refinements, the time and cost of development could be largely reduced if certain aspects of the user interface were ‘standard equipment.’ [Frakes, 1994, Lim, 1994] “Traditional ITSs are concentrated on the domain knowledge they are supposed to present and teach; hence their control mechanisms are often domain-dependent” [Jericin and Devedzic, 2000]. More recent ITSs pay more attention to generic problems and concepts of the tutoring process. They try to separate architectural, methodological, and control issues from domain knowledge as much as possible. This is due to the improved interactive capability of new technologies, it is now possible to create learning environments on computers that are capable of providing students with feedback, managing large on-screen graphical simulations, provide on-screen enrichment, and in-flight assessment [Barron *et. al.*, 1998; Greenfield and Cook, 1996; Scardamalia, 1993; Hmelo and Williams, in-press; Kafai, 1995; Linn *et. al.*, 1996]. Clearly, ITS’s are a rich area of investigation; however, the diffuse needs involved in building specific expertise systems has limited the focus on helping build general learning intelligence.

We foresee that the mechanism by which the transient emotional state of the user is recognized and appropriately responded to could be separated from specific ITSs and made into a stand-alone system, as is the aim with our proposed Learning Companion. This module might be a “plug-in” to the “standard” Tutor/Pedagogical module, or might stand alone, beside the system – appearing to the learner to be on his or her side, vs. on the side of the “expert” tutor.

In general ITS modules (and models) can be fairly powerful,. For example, students using the LISP tutor [Anderson, 1990] completed programming exercises in 30% less time than those receiving traditional classroom instruction and scored 43% higher on the final exam. However ITSs “were not designed with significant user input and do not address the practical issues encountered when educators actually use these systems,” [Jericin and Devedzic 1997]. In other words they were designed primarily to be a knowledge base much the same as a CD-ROM encyclopedia rather than as an assistive tutor. Our focus differs from that of ITS developers in taking on the role of a *learning companion that helps teach learning intelligence*—a collaborator with the student, there to help the student learn, and in so doing, learn how to learn better—not a database with a teaching pedagogy.

We agree with those who think that ITSs need to learn not only to recognize affective expressions but also to intelligently adapt their behavior based upon this information. This involves continuous learning with user feedback as well as skillful communication and management of emotion information. It is important to concentrate equally on the design of cognitive and affective aspects—learning and teaching issues—as on the design of the Domain Knowledge module. We believe that our Learning Companion could prove to be a reliable plug-in module for ITSs, as well as a useful stand-alone application. We expect that eventually there may be learning companions in many different styles and “personalities”, so that users can have a choice. Tutors and companions might also adopt complementary styles. We also expect that future companions might, in many cases, have personalized learning algorithms, allowing them to adapt their strategy to individual students based on longer-term observations of interactions with that learner, thus customizing their feedback for maximal success, similar to how experienced mentors adapt a large repertoire of strategies to better help an individual. But, these are areas for future work; first, we proposed to build a basic system that will allow investigation of the usefulness of a computerized learning companion.

Key questions

A major difficulty with developing ITSs or an Affective Learning System (e.g., our *Learning Companion*) is that a monolithic effort seems to focus on making the computer system smart, both about the topic and about pedagogy as opposed to helping the student learn to learn. In contrast, the Learning Companion promises to be a separate module, usable with or without an ITS, focused solely on helping the learner to learn. A Learning Companion’s expertise is not subject-dependent, although we do expect to tailor it to facilitate a model-based reasoning style, which we believe is key for SMET learning, in contrast with a “memorize the facts and rules” style, which is very limited. The Learning Companion’s expertise will be targeted at helping the learner with meta-cognitive and affective tasks, such as “Is learning on-track or off?, and “Is the learner getting frustrated and in need of a change of pace or feedback?” Although initially we will aim for simple discrimination of signs of on-goal/off-goal learning, we recognize that the mood set by the companion can also influence the learner, and thus we may explore different styles, “moods” and personalities for infecting the learner with zeal appropriate to the learning experience.

Whether a module of an ITS or a stand-alone system, the *Learning Companion* aims to: sense affective and cognitive aspects of the learning experience, help the learner better understand these aspects, and work with the learner, responding to his/her state, to help him/her improve not just knowledge, but mastery of the learning process. The proposed system will facilitate:

- Observing and trying to understand the processes a learner experiences during an *off-goal* episode (e.g., cognitive assessment, cognitive appraisal, one emotion or another), which will assist in bringing a learner back *on-goal*,
- Attending to the affective state of a learner and crafting responds in an appropriate/intelligent manner, in order to improve the quality of learning,
- Adjusting the learning environment in response to the transient emotions of the learner, for example, help the learner ‘see’ that their frustration is associated with an over-complex task, and help them break the task into simpler elements, and,
- Understanding how to spawn the most critical emotions in order to facilitate maximal results, which, for example, can spawn the *teachable moment*.

VI. Methods of Investigation

Do emotions contribute to intelligence, and if so, what are the implications for the development of a technology of affective computing?

- Robert Provine, *What Questions Are On Psychologist's Minds Today?*

We propose to explore, describe, and evaluate ways to facilitate the cognitive and affective learning processes inherent for SMET learning. We intend to incorporate these research-based findings into a testbed that we will design and build—the *Learning Companion* (a software-based interactive application), which may be plugged-into several commercially available software packages or into some home-grown simulations. The intent is to recognize aspects of the affective/cognitive state of the learner and respond in a manner that will optimally improve the learning journey.

Specifically...

We propose to develop a prototype automated Learning Companion. The Learning Companion will react to a subject in-need, i.e., when they show signs of becoming confused, distracted, anxious, worried, etc. The Learning Companion prototype will assist the learner when necessary—specifically, when the learner requests help, becomes ‘stuck,’ or makes a mistake serious enough to warrant a response. The Learning Companion will also allow systematic exploration for comparing different kinds of interventions, different affective triggers of interventions, and different styles and means of presenting the interventions.

Our mission is to develop and test the feasibility of using our software-based automated *Learning Companion* to support, encourage, and guide a learner through a learning journey. We will extract pedagogically relevant information that might be of use to diagnose and remediate learning, and refine the *Learning Companion*. We will embody the *Learning Companion* with behaviors modeled after expert SMET teachers. While the learner is using the prototype simulation the human coach (who will be hidden-away as was Professor who controlled the *Wizard of Oz* from behind a curtain) will only intervene if the learner becomes stuck, is operating on inaccurate assumptions, etc. We expect the *Learning Companion* will provide assistive comments, hints, or guidance in a patient and supportive manner using predigitized recordings of a human speaker, or using silent text that would appear on-screen.

Our prototypes (initially simulated and increasingly automatic) will be evolved and tested on Massachusetts’s school children and will employ the use of several state-of-the-art-sensing technologies in authentic learning activities: on-screen buttons, a sensing mouse, and perhaps other pattern recognition devices that have been built at the MIT Media Lab and other institutions. The aim of using these technologies will be to assess what affective information a computer can accurately and comfortably obtain without adversely interfering with the learning process. We will also use expert human observations to assess the accuracy of the technology-based ones. Our initial goal is to, at least, describe where in the emotion axis space the student might be. We will also look for ways to refine this description based on other situational and behavioral variables (such as repeated banging on the mouse, which is more likely to be a sign of frustration than a sign of dispiritedness.)

Based upon our model (Figures 1, 2a, 2b) we will hypothesize intervention strategies that could be aligned with the model (hints, encouragement, etc.) and test these against controls (e.g., timed interventions) to see if the model-based interventions outperform the controls. We will perform a factor analysis of a liberal list of emotions with the intent of identifying the most relevant ones and discarding the ones with negligible impact. The relevant emotions will be infused into the *Learning Companion*, which will be field-tested on a sample population of learners. This phase may initially be implemented with a person in the loop, depending

on the success with phase two, but our goal will always be toward automating the intervention strategy. This phase will also help us test the success of the model, and from those findings we will make refinements that better fit the experience of real SMET learners.

We will evaluate three aspects of the development of the Learning Companion before moving to the next; effectiveness of the simulated Learning Companion, accuracy in detecting a need for intervention, and effectiveness of the automated Learning Companion:

- Effectiveness of the simulated Learning Companion

To facilitate quicker measurement of the educational effectiveness of the Learning Companion's interventions we propose to utilize a "Wizard of Oz" approach; the simulated Learning Companion appears to be automatic to the subjects, but is actually controlled by an out-of-sight human experimenter. The experimenter will listen to and watch the subjects as they proceed through the software-based simulation (The Incredible Machine, Gizmos and Gadgets). The 20 subjects will be selected from 3 Massachusetts public schools. The subjects will be exposed to two software simulations with accompanying comprehension questions. We will perform a within-subject controlled experiment. All subjects will be exposed to the two software programs. In the experimental condition, the students will be presented the software with the (simulated) Learning Companion. In the control condition, the student will be exposed to the other software package without the assistance of the Learning Companion. Comprehension questions will be administered to this group after the completion of the trial. The experimental design will be counterbalanced both by the software package that each student received (Learning Companion assisted and not assisted), and by the order of conditions. We also plan to design a post-study evaluation (via either questionnaire or behavior) in order to provide *proof-of-concept* that the subject is better-off.

- Accuracy in detecting a need to intervene—Automating the *Learning Companion*

The second experiment will measure the *Learning Companion's* ability to automatically detect occasions where it should intervene. The purpose of this experiment will be to obtain data that could be used to improve the ability of the *Learning Companion* to recognize 'when' and 'how' to intervene. We expect to measure the ability of the automated *Learning Companion* to appropriately intervene. In this experiment we will play the video recordings of automated *Learning Companion's* activity to several expert teachers and they will evaluate the timeliness and appropriateness of the *Learning Companion's* actions.

- Effectiveness of the automated *Learning Companion*

The third experiment will test the overall effectiveness of the *Learning Companion*. Its purpose will be to identify limitations of the *Learning Companion*, and how those limitations might be overcome. The subjects for this experiment will be 30 students taken from cooperating Massachusetts public schools spanning a spectrum of academic abilities. The material for this experiment will be similar to those in step just above. To reduce the time required per subject, we plan to use a more streamlined within-subjects design. Each subject will be tested in two conditions. In the experimental condition, the subject will run the software with the assistance of the *Learning Companion*. In the control condition the subject will run the software without the assistance of the *Learning Companion*. Choice of software program and order of condition will counterbalance this experiment design. To reduce start-up effects, the subjects will be trained in each condition. The dependent measure will consist of the score from the comprehensive assessment tool. The subjects' scores from a pre-test will measure learning intelligence; we propose to evaluate existing tools and explore the feasibility of designing a tool of our own to accomplish this task.

We plan to characterize which subjects benefited most from the various treatments. Although it is hard to make such claims, we may, for example, plot individual effect size (between-condition difference in comprehension score) versus overall SMET score on the MCAS. We may also look at independent teacher's reports of student motivation about learning, and we may look at time spent on tasks that are not required (e.g., how long the learner chooses to engage in a learning experience beyond the required time.) We expect behavioral measures to come closer to measuring benefits of the system than would simple test results, although both kinds of measures can be illuminating.

However, the primary value of this experiment will be to identify the most important modifications required to improve the Learning Companion's usability, robustness, and effectiveness. For this reason, we plan to collect two types of data. First, video recordings will capture all subject-Learning Companion interaction. Second, the 'event files' will record Learning Companion's date-stamped inputs (e.g., mouse clicks, affective information sensed from user, Learning Companion actions).

To analyze the data, we plan to, first, review the videotape to identify problematic interactions. Then we plan to use the event files to characterize the problems. We expect to find many design improvements as a result in several areas (e.g., recognition error, time lag, and interface design).

Recognition error—The ‘Wizard of Oz’ experiments used to develop the interface will use a human operator to identify ‘when’ the Learning Companion should assist/intervene. We expect that, in general, and certainly, in the initial prototypes the Learning Companion will be less accurate and effective in automated mode as compared to operator-controlled (Wizard of Oz) mode. We expect that recognition errors will occur in two categories: false alarms and undetected miscues. ‘False alarms’ occur when the Learning Companion assists/intervenes when it should not. ‘Undetected miscues’ occur when the Learning Companion does not assist/intervene when it should.

Time lag—A second source of problem is the time lag between user’s input and the Learning Companion’s response. We envision problems such as: a student hesitates for a few seconds and the Learning Companion views this hesitation as a need for assistance, the student then continues on, only to be interrupted by the Learning Companion.

Interface confusion—Although we expect the children to be comfortable with the interface almost immediately, occasionally the interface may confuse them. For example, it may be unclear how to manually activate Learning Companion assistance.

We will work to craft solutions to these foreseen (and future unforeseen) problems; it is crucial that the Learning Companion not be perceived as irritating or annoying, and when it can’t do the right thing, it may be best for it to do nothing at all, or to at least degrade softly with an acknowledgment of its limitations.

VII. Conclusion

Evaluation and Dissemination

In disseminating our results (in the form of computational methods, implemented code, data, and analyses), we plan to build upon strong records of dissemination through conferences and published papers.

Our strategy for the proposed research is to deploy new functionality on a experimental basis as early as possible. This strategy will provide early value to the participating schools for their students, and early feedback to us on usability and effectiveness, letting educational needs guide our research.

Therefore we propose to aim for pilot development of the implementation portion of the *Learning Companion* at 2-3 Massachusetts public elementary schools in year 2. This will be followed by experimental deployment in one or more additional Massachusetts public schools. Deployment will start in classrooms of early-adopter teachers who understand the experimental nature of the software and are willing to help shape it. This schedule and the timetable of subsequent action will depend on the availability of equipment, personnel, and other separately funded resources for experimental deployment.

Success of the experimental deployment in Massachusetts will pave the way for wider dissemination and technology transfer into affordable commercial products for the school and home market.

The proposed methods for analysis of SMET ability and for development of the *Learning Companion* will be extended beyond improved miscue detection to identify specific miscues, exploit temporal clues, and recognize pedagogically significant patterns.

References

- Anderson, J., (1990). Analysis of Student Performance with the LISP Tutor. In Frederiksen, N., Glaser, R., Lesgold, A.M. and Shafto, M. (Eds.), Diagnostic Monitoring of Skill and Knowledge Acquisition, Hillsdale, NJ: Lawrence Erlbaum.
- Azmitia, M. (1988). Peer interaction and problem solving: When are two heads better than one? Child Development, 59, 87-96.
- Barron, B.J., et. al. (1995). Doing with understanding: Lessons from research on problem-solving and project based learning, Journal of Learning Sciences.
- Bartlett, M., J.C. Hager, P. Ekman, and T.J. Sejnowski, (1999). "Measuring facial expressions by computer image analysis," Psychophysiology, vol. 36, pp.253--263.
- Bonar, T., (1988). Byte-sized Tutor. In J. J. Psozka, L. D. Massey and S. A. Mutter (Eds.) Intelligent Tutoring Systems: Lesson Learned, Lawrence Erlbaum Associates.
- Briggs, P. (1996). Self confidence as an issue for user-modeling. In Proceedings of the Fifth International Conference on User Modeling, Kailua-Kona, Hawaii.
- Brusilovsky, P., S. Ritter and E. Schwarz, (1997). Distributed Intelligent Tutoring on the Web. In du Boulay, B., Mizoguchi, R. (Eds.) Artificial Intelligence in Education, IOS Press, Amsterdam OHM Ohmsha, Tokyo, 482-489.
- Chen, L.S., T.S. Huang, T. Miyasato, and R. Nakatsu, "Multimodal human emotion/expression recognition," in Proc. of Int. Conf. on Automatic Face and Gesture Recognition, (Nara, Japan), IEEE Computer Soc., April 1998.
- Cohn, J.F., A. J. Zlochower, J. Lien, and T. Kanade, Automated face analysis by feature point tracking has high concurrent validity with manual FACS coding, Psychophysiology, vol. 36, pp.35-43, 1999.
- Collins, A. (1996) Design Issues for Learning Environments, in S. Vosniadou, E. Decourt, et.al., (Eds.) International Perspectives on the Psychological Foundations of Technology-based Learning Environments, Lawrence Erlbaum Assoc. Hillsdale, NJ.
- Csikszentmihalyi, M. (1990). Flow: The Psychology of Optimal Experience, Harper-Row: NY.
- Damasio, A.R., (1994). Descartes Error: Emotion, Reason and the Human Brain, G.P. Putnam Sons: NY.
- del Soldato, T. (1994). Motivation in Tutoring Systems. Tech. Rep. CSRP 303, School of Cognitive and Computing Science, The University of Sussex, UK.
- DeSilva, L.C., T. Miyasato, and R. Nakatsu, Facial emotion recognition using multi-modal information, in Proc. IEEE Int. Conf. on Info., Comm. and Sig. Proc., (Singapore), pp. 397-401, Sept 1997.
- Dertouzos, Michael. The People's Computer: Speech and Vision, MIT Technology Review, May-June 2000, p. 24.
- de Vincente, A. and H. Pain. (1998). Motivation Diagnosis in Intelligent Tutoring Systems.
- Doise, W. , Mugny, G., and Perret-Clermont, A. (1975). Social interaction and the development of cognitive operations. European Journal of Social Psychology, 5(3), 367-383.
- Doise, W., Mugny, G., & Perret-Clermont, A. (1976). Social interaction and cognitive development: Further Evidence. European Journal of Social Psychology, 6, 245-247.

- Donato, G., M.S. Bartlett, J.C. Hager, P.Ekman, and T.J. Sejnowski, Classifying facial actions, IEEE T. Pattern Analy. and Mach. Intell., vol. 21, pp. 974--989, October 1999.
- Ekman, Paul. (1992). Are there basic emotions?, Psychological Review, 99(3):550-553.
- Ekman, P., (1997). Facial Action Coding System, Consulting Psychologists Press.
- Ellis, S., Klahr, D., and Seigler, R.S. (1993). Effects of feedback and collaboration on changes in children's use of mathematical rules. A paper presented in Society for Research in Child Development. New Orleans.
- Essa I. and A. Pentland, Coding, analysis, interpretation and recognition of facial expressions, IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 19, pp. 757--763, July 1997.
- Essa, I., Analysis, Interpretation and Synthesis of Facial Expressions. Ph.D. thesis, MIT Media Lab, Cambridge, MA, Feb. 1995.
- Fischer, G., (1993). Shared Knowledge in Cooperative Problem-Solving Systems – Integrating Adaptive and Adaptable Components, in M. Schneider-Hufschmidt, T. Kuhme, *et.al.* (Eds), Adaptive User Interfaces: Principles and Practice, North Holland, Elsevier Science Publishes B.V., Amsterdam.
- Frakes W. B., (1994). Success Factors of Systematic Reuse. IEEE Software 11(5): 15-19.
- Gabbert, B., Johnson, D. W., and Johnson, R.T. (1986). Cooperative learning, group-to individual transfer, process gain, and the acquisition of cognitive reasoning strategies. The Journal of Psychology, 120(3), 265-278
- Gardner, R.C., (1985). Social psychology and second language learning: The role of attitudes and motivation. London: Edward Arnold.
- Goleman, D., (1995). Emotional Intelligence. Bantam Books: New York.
- Guzdial M., (1995). Software Related Scaffolding to Facilitate Programming in Science Learning. Interactive Learning Environments, 4(1), 1-44.
- Hansen, J (1999).Communication during ICASSP '99 panel on Speech Under Stress.
- Healey, J. (2000). Wearable and Automotive Systems for Affect Recognition from Physiology. Ph.D. thesis, MIT Media Lab
- Hioe, W. and D.J. Campbell (1988). An expert system for the diagnosis of motivation-related job performance problems: Initial efforts. DICS publications no. TRA3/88, Department of Information Systems and Computer Science, National University of Singapore.
- Hmelo, C. and S.M. Williams (in-press). Special Issue: Learning through problem-solving. The Journal of Learning Sciences 7(3 and 4).
- Huang, T.S., L.S. Chen, and H. Tao, Bimodal emotion recognition by man and machine, ATR Workshop on Virtual Communication Environments, (Kyoto, Japan), April 1998.
- Holt, John (1967) How Children Learn, Dell Publishing Co.

- Isen, Alice, M. (1999) Positive Affect and Decision Making, in Michael Lewis and Jeanette M. Haviland, eds., Handbook of Emotions, 2nd ed., The Guilford Press.
- Issroff, K. and del Soldato, T., Incorporating motivation into computer-supported collaborative learning. In Brna, P., *et. al.* (Eds.), Proceedings of the European Conference on Artificial Intelligence in Education, 1996.
- ITS Working Group. (1995a). Working Group: ITS Shells and Generic Task Domains [On-line]. Available at : <http://advlearn.lrdc.pitt.edu/its-arch/group-shells.html> as of May 10, 2000.
- ITS Working Group. (1995b). Working Group: Exploring Industry Standards [On-line]. Available at : <http://advlearn.lrdc.pitt.edu/its-arch/group-cots.html> as of May 10, 2000.
- ITS Working Group. (1995c). Working Group: Communication Architectures for ITS Components [On-line]. Available at : <http://advlearn.lrdc.pitt.edu/its-arch/group-commarch.html> as of May 10, 2000.
- Jackson, Shari L., Joseph Krajcik, Elliot Soloway (1999). The Design of Guided Learner-Adaptable Scaffolding in Interactive Learning Environments, Available at: http://hi-ce.eecs.umich.edu/papers/design_of_guided/index.html
- Jerinic, L. and Devedzic, V., OBOA Model of Explanation Module in Intelligent Tutoring Shell. SIGCSE Bulletin, Vol. 29, Number 3, September 1997, ACM PRESS, 133-135.
- Jerinic, L. and Vladan Devedzic (2000). The Friendly Intelligent Tutoring Environment: Teacher's Approach, SIGCHI Bulletin, Vol. 32 No. 1. p. 83.
- Joiner, R. Littleton, K. and Riley, S. (1991) Peer Presence and Peer Interaction in Computer Based Learning. Paper presented to the BPS Developmental Psychology Section Annual Conference, University of Cambridge, September 1991.
- Kafai, Y.B. (1995). Minds in Play: Computer Game Design as a Context for Children's Learning. Hillsdale, NJ: Erlbaum.
- Klein, J., Y. Moon and R. W. Picard (Feb., 1999), This Computer Responds to User Frustration, Proceedings of CHI.
- Klein, J., Computer Response to User Frustration, (1999b)MIT Master's degree thesis.
- Keller, J.M. (1983). "Motivational design of instruction. In C.M. Reigeluth (Ed.). Instructional design theories and models: An overview of their current status." Hillsdale, NJ: Erlbaum.
- Keller, J.M. (1987a, Oct.). Strategies for stimulating the motivation to learn. Performance and Instruction, 26(8), 1-7. (EJ 362 632).
- Keller, J.M. (1987b). IMMS: Instructional Materials Motivation Survey. Florida State University.
- Keller, J.M. & Keller, B.H. (1989). Motivational Delivery Checklist. Florida State University.
- Kort, B., and R. Reilly (2000). Detecting and Interpreting Emotions of Students: Defining the First Alert, MIT Media Lab Technical Report.
- LeDoux J.E., Emotion, Memory and the Brain. Scientific American, June 1994, pp. 50-57.
- Leibniz, G.W.V., Monadology and Other Philosophical Essays. Indianapolis: The Bobbs-Merrif Company, Inc., 1965. Essay: Critical Remarks Concerning the General Part of Descartes' Principles (1692), Translated by: P. Schrecker and A. M. Schrecker.

- Leidelmeijer, K., Emotions: An Experimental Approach. Tilburg University Press, 1991.
- Lepper, M.R. and R.W. Chabay (1988). Socializing the intelligent tutor: Bringing empathy to computer tutors. In Heinz Mandl and Alan Lesgold (Eds.), Learning Issues for Intelligent Tutoring Systems, pp. 242-257.
- Lessig, L., (1999). Code and Other Laws of Cyberspace, Basic Books, NY.
- Lewis M., (1993). Ch. 16: The emergence of human emotions. In M. Lewis and J. Haviland, (Eds.), Handbook of Emotions, pages 223-235, New York, NY. Guilford Press.
- Lim W. C., (1995). Effects of Reuse on Quality, Productivity and Economics. IEEE Software 11(5): 23-29.
- Linn, M.C., N.B. Songer, and B.S. Eylon. (1996). Shifts and convergences in science learning and instruction, in R.C. Calfee and D.C. Berliner (Eds.) Handbook of Educational Psychology. Riverside, NJ: MacMillian.
- Loveridge, N. Joiner, R., Messer, D., Light, P. and Littleton, K. (1993). Social Conditions and Computer Based Problem Solving. Paper presented at the fifth meeting of the European Association for research into learning and instruction, Aix en Provence, France, August 31.
- Matsubara, Y. and Nagamachi, M., (1996). Motivation systems and motivation models for intelligent tutoring, in Claude Frasson, et. al., (Eds.), Proceedings of the Third International Conference in Intelligent Tutoring Systems, pp. 139-147.
- Mayer, John D., (1986). How Mood Influences Cognition, in N.E. Sharkey (Ed.) Advances in Cognitive Science (pp. 290-314). Chichester, West Sussex: Ellis Horwood Ltd.
- Mercer, N (1994). The Quality of Talk in Children's Joint Activity at the computer. Journal of Computer Assisted Learning (10):24-32.
- Mostow, Jack D, (1996). An Automated Reading Assistant That Listens. NSF MDR Grant #9616546
- Myers, David G, (1998). What Questions Are On Psychologist's Minds Today? Available on-line as of May 1, 2000 at: http://www.edge.org/3rd_culture/myers/index.html
- Nathan, J., et.al. (1990). A Theory of Algebra Word Problem Comprehension and Its Implications for Unintelligent Tutoring Systems, (Technical Report 90-02) Institute of Cognitive Science, University of Colorado, Boulder.
- Niedenthal, P. M. (1990). Implicit perception of affective information, Journal of Experimental Social Psychology, Vol. 26(6), 505-527.
- Olson, R.K. and Wise B. (1987). Computer Speech in Reading Instruction, In D Reinking (Ed.). Computers and Reading: Issues in Theory and Practice. New York: Teachers College Press.
- Papert, Seymour (1993). The Children's Machine: Rethinking School in the Age of the Computer, Basic Books: New York.
- Picard, R.W., (1997). Affective Computing. Cambridge, Mass.: MIT Press 1997.
- Plutchik, R. 'A general psychoevolutionary theory of emotion,' in Emotion Theory, Research, and Experience (R. Plutchik and H. Kellerman, eds.), vol. 1, Theories of Emotion, Academic Press, 1980.
- Porter, L.W. & Lawler, E.E. (1968). Managerial attitudes and performance. Homewood , : rsey Press.

- Provine, Robert, (1998). What Questions Are On Psychologist's Minds Today? Available on-line as of May 1, 2000 at: http://www.edge.org/3rd_culture/myers/index.html
- Pstotka, J. J., Massey, L. D. and Mutter, S. A. (Eds.) Intelligent Tutoring Systems: Lesson Learned, Lawrence Erlbaum Associates, 1988.
- Reeves, B., and Nass C.I., (1996). The Media Equation: How People Treat Computers, Television, and New Media Like Real People and Places. Cambridge University Press.
- Resnick, M. A Media Lab for Kids: \$27 Million from Isao Okawa Creates Center for Future Children at MIT, MIT News, November 18, 1998.
- Reilly, R., (1999). Policy Schmolicy—It's the Architecture Stupid!, MultiMedia School, November-December 1999.
- Russell, D., Moran T. P., Jordan, D. S., (1988). The instructional design environment. In J. J. Pstotka, L. D. Massey and S. A. Mutter (Eds.) Intelligent Tutoring Systems: Lesson Learned, Lawrence Erlbaum Associates.
- Scardamalia, M., and C. Bereiter. (1993). Technologies for knowledge-building discourse. Communications of the ACM 36(5): 37-41.
- Scheirer, J., R. Fernandez and R. W. Picard (1999), Expression Glasses: A Wearable Device for Facial Expression Recognition, Proceedings of CHI, February, 1999.
- Sharan, S. (1980). Cooperative learning in small groups: Recent methods and effects on achievement, attitudes, and ethnic relations. Review of Educational Research, 50 (2), 241-271.
- Sleeman, D., (1987). PIXIE: A Shell for Developing Intelligent Tutoring Systems. Artificial Intelligence in Education, Vol. 1, 239-265.
- Soloway, E., Guzdial, M., & Hay, K. E. (1994) Learner-Centered Design: The Challenge for HCI in the 21st Century, Interactions, Vol. 1, No. 2, April, 36-48.
- Soloway, E., Jackson, S. L., Klein, J., Quintana, C., Reed, J., Spitulnik, J., Stratford, S. J., Studer, S., Eng, J., & Scala, N. (1996) Learning Theory in Practice: Case Studies of Learner-Centered Design. In ACM CHI '96 Human Factors in Computer Systems, Vancouver.
- Soloway, E. (1999). Scaffolded Technology Tools to Promote Teaching and Learning in Science [On-line]. Available as of May 22, 2000
- Teasley, S.D. (1992). Communication and collaboration: The role of talk in children's peer collaboration. Doctoral dissertation, University of Pittsburgh.
- Tudge, J. (1989). When collaboration leads to regression: Some negative consequences of socio-cognitive conflict. European Journal of Social Psychology, 19, 123-138.
- Wolf, B., (1987). Theoretical Frontiers in Building a Machine Tutor. In Kearsley, G. P. (Ed.) Artificial Intelligence and Instruction-Application and Methods, Addison-Wesley, Reading, 229-267.
- Yacoob Y. and L. Davis, "Recognizing human facial expressions from log image sequences using optical flow," IEEE T. Pattern. Analy. and Mach. Intell., vol. 18, pp.636--642, June 1996.
- Vail, Patricia, (1994). Emotions: The On/Off Switch for Learning, Modern Learning Press.
- Vroom, V.H. (1964). Work and motivation. New York: Wiley.

- Vygotsky, L., (1978). Mind in Society. Cambridge, Mass: Cambridge University Press.
- Vyzas, E. and Picard. R. (1999). Online and Offline Recognition of Emotional Expression from Physiological Data, Workshop on Emotion-based Agent Architectures, International Conference on Autonomous Agents, Seattle, WA.
- Wegerif, R. (1996). Using Computers to Help Coach Exploratory Talk Across the Curriculum. Computers in Education 26 (1-3):51-60. ISSN-0360-1315.
- Whitelock, D. Taylor, J., O'Shea, T. (1993). What do you say after you have said hello? Dialogue analysis of conflict and cooperation in a computer supported collaborative learning environment. PEG 93 Conference, Edinburgh Scotland, August 1993.
- Whitelock, D. and Scanlon, E., (1996). Motivation, media and motion: Reviewing a computer supported collaborative learning experience. In Brna, P., et. al. (Eds.), Proceedings of the European Conference on Artificial Intelligence in Education, 1996.
- Wood, D., Bruner, J.S. and Ross, G. (1975). The Role of Tutoring in Problem Solving, Journal of Child Psychology and Psychiatry, 17, 89-100.
- Youngblut, C. (1995). Government Sponsored Research and Development Efforts in the Area of Intelligent Tutoring Systems: Summary Report Inst. For Defense Analyses Paper No. P-3058.
- Zajonc, R.B., (1998). Emotions, In Gilbert D.T., Fiske, S.T. and Lindzey, G. (Eds.) The Handbook of Social Psychology, Vol. 1 (4th Edition), pp. 591-632.