Project Description

Vision

The last half-century of technological acceleration has yielded a massive incursion of digital technology into the learning environment, making dramatic differences, and promising even greater differences, to the practice of learning. Computers have served as tools to aid in learning at all levels from simple classroom activities, to the way theorists think about thinking. The field of artificial intelligence, with emphasis on ideas such as knowledge representation, modeling of logical processes, and other kinds of important cognitive activities, has prompted thinking about parallel concepts in human learning, and facilitated the development of theories where thinking and learning are viewed as information processing. Both human and machine learning research have benefited from this exchange of ideas surrounding cognition and computation.

In the last decade, a variety of findings in neuroscience, psychology, and cognitive science have supported a different view of cognition: one where affect is complexly intertwined with thinking, and performs important functions with respect to guiding rational behavior, memory retrieval, decision-making, creativity, and more. While it has always been understood that too much emotion is bad for rational thinking, recent findings suggest that so too is too little emotion: when basic mechanisms of emotion are missing in the brain, then intelligent functioning is hindered. These findings point to new advances in understanding the human brain not as a purely cognitive information processing system, but as a system in which affective functions and cognitive ones are inextricably integrated with one another

Most teachers and theorists know that affect plays a crucial role, recognizing it under poorly understood headings like motivation, emotion, interest, and attention. However, scientific understanding of affect in learning, and tools that transport this into practice, are still in their infancy. Our vision is to fundamentally advance the understanding of affect in learning through development and use of technology. Our goal is to redress the imbalance between affect and cognition, and between theory and practice, to bring balance to the science of learning and to its technologies.

We plan to carry out this vision by (1) building tools and technologies that elicit, sense, communicate, measure, and respond appropriately to affective factors; (2) building new models that incorporate affect, as a foundation for both new approaches to education and more effective machine learning; and (3) building tools and environments that foster new levels of both creative expression and conceptual understanding by more effectively engaging people's senses and bodies and leveraging their interests and passions. Additionally, and interwoven with these technology construction efforts, we plan to develop and refine theories and terminology related to affect in learning.

MIT is ideally situated to host an NSF Science of Learning Center. MIT has a long tradition of emphasizing itself not as a research institute or a teaching institute, but as a learning institute, where undergraduates, graduates, faculty, and others collaborate in an ongoing process of discovery. MIT pioneered an undergraduate research opportunities program, which has been adopted at many other institutions. Recently, the MIT faculty announced the publication of all MIT course materials, free and open to the world (http://ocw.mit.edu). This OpenCourseWare initiative, which will have the full catalog online by 2007, has already had a significant impact internationally and domestically. Following in this spirit, we propose to develop an NSF Science of Learning Center at MIT that does not limit partnerships to just a few selected institutions, but rather opens up what we do to the world, through a variety of new initiatives.

To some, it might seem unusual for a technology leader like the MIT Media Lab to focus on "affect," since many people associate technology (and, particularly, information technology) with cognitivist and rationalist perspectives; nonetheless, the Media Lab, which would form the hub of the proposed Center for Affective Learning, has always been in the forefront of taking a fuller approach to technology. Since its inception, it has gathered faculty from multiple disciplines such as psychology, mathematics, cognitive science, physics, computer science, engineering, and more into an open intellectual environment that emphasizes both human and machine, both arts and technology, and both theory and practice. The Media Lab currently has a unique combination of researchers leading the world in efforts that develop technologies that sense and respond to affect, learning systems and theories that involve modeling attention, motivation, and other affective phenomena, and tools and environments for facilitating new kinds of learning and expression. Our researchers are united in their interest in the roles of emotion, motivation, attention, caring, and other affective phenomena in learning, and are eager to collaborate on projects advancing the state of the art in both human and machine learning. The development of an NSF Center for Affective Learning would allow these efforts to advance in bold new ways, to take significantly greater scope, and to fundamentally impact the science of learning.

Challenges in Affective Learning

Scientific findings over the past decade have started to lay the foundation for a better understanding of the role of affect in learning. Research has demonstrated, for example, that a slight positive mood doesn't just make you feel a little better but also induces a different kind of thinking, characterized by a tendency toward greater creativity and flexibility in problem solving, as well as more efficiency and thoroughness in decision making. These effects have been found among many groups of different ages and professions (Isen, 2000). Not only positive mood, but also affective states such as fear, anger, sadness, and joy show up in the brain as different patterns of blood flow, providing one possible explanation for how affect influences brain activity (e.g., Lane et al., 1997; Damasio et al., 2000). In the case of positive affect, a theory of two separate but interacting dopamine systems has been proposed for mediating some of the effects positive affect has on cognition (Ashby et al., 1999). There is also some indication that positive affect increases intrinsic motivation (Estrada et al., 1994). Although the work in this area is only beginning to be launched, it already suggests that a positive mood is not best for all kinds of thinking, but that certain affective states facilitate some kinds of thinking better than others. Learning research has long recognized the importance of facilitating different ways of thinking—with beliefs such as "you don't understand something unless you understand it in many ways." In his forthcoming book, *The Emotion Machine*, Marvin Minsky argues, "…when we change what we call our 'emotional states,' we're switching between different *ways to think*." (Minsky, 2003).

Among educators and educational researchers, there is a growing recognition that interest and active participation are important factors in the learning process (e.g., Bransford et al., 1999). But acceptance of these ideas is based largely on intuition and generalized references to constructivist theorists (Piaget, 1960; Vygotsky, 1962; Vygotsky, 1978). There is need for new types of studies on the role of affect in learning. We believe that new technologies can play a particularly important role in these efforts, helping us to measure, model, study, and support the affective dimension of learning in ways that were not previously possible.

Terminology and Theories

One of the problems with studying affect is defining what it is—illuminating a better definition of affect and related terms like emotion, motivation, caring, and so forth. Modern research in this area began before the turn of the last century, when Charles Darwin (1872) and William James (1890) devoted seminal works to describing emotion, anchoring its description in measurable bodily changes and expressions. In the last century many cognitive scientists and psychologists have advanced theories and definitions of emotion, motivation, and other affective phenomena (e.g., Tomkins, 1962a,b; Clynes, 1977; Buck, 1984; Kagan, 1984; Collier, 1985; Frijda, 1986; Ortony et al., 1988; Plutchik and Kellerman, 1980-90; Lazarus, 1991; Oatley, 1992, and more). Nearly a hundred definitions of emotion had been categorized as of 1981 (Kleinginna and Kleinginna, 1981) when Don Norman wrote his now classic essay naming emotion as one of the twelve major challenges for cognitive science (Norman, 1981).

Today, the burgeoning literature on affect includes diverse communities such as psychology, cognitive science, neuroscience, engineering, computer science, philosophy, and medicine, and this has contributed to a similarly diverse understanding of numerous basic terms related to affect—such as "emotion," "motivation," "attention," "reward," and more. One of the challenges we will address in the Center for Affective Learning is to bring together top theorists and practitioners from different fields in order to refine the language used with respect to affect and learning. The emphasis will be on combining ideas from people who analyze affective phenomena with ideas from people who try to elicit or synthesize them—to try to bring these two ends together, and to insure that the language is rooted in conceptual theories, human and animal observation, and in practical understanding of what elements "make a learning system work." These gatherings will also have as an aim, the introducing of researchers who approach the same problems from different vantage points, in order to facilitate their collaboration and break down some of the disciplinary boundaries that have held back progress. We thus see these forums not only as producing a more workable language for researchers, but also as contributing to the gargantuan task of beginning to unify different levels of thinking about these phenomena—from the high levels of human behavior, to the lower levels of synaptic chemistry. As part of the plans for the "hub" nature of the Center, these conversations will be web-cast, archived, and made searchable, and in some cases published in additional forums.

Although there are dozens of books on various affective phenomena, there is a lack of theories that engage the topic of affect in learning. Some of the classic works on affect emphasize cognitive and information processing aspects in a way that can be encoded into machine-based rules, and studied in a learning interaction. The most widely adopted of these is the OCC model of emotion (Ortony et al., 1988); however, this model does not include many of the affective phenomena observed in natural learning situations (such as interest, boredom, or surprise). Csikszentmihályi (1991) has emphasized the tendency for a pleasurable state of "flow" to accompany problem solving that is neither too easy nor too challenging, and there have been other scattered attempts to address emotions involved in learning (e.g., Lepper and Chabay, 1988; Mandler, 1984; Kort et al., 2001a,b). However, there is still very little understanding as to which emotions are most important in learning, and how they influence learning. To date there is no comprehensive, empirically validated, theory of emotion that addresses learning. With respect to motivation in learning, there has been much more work and much more progress, illuminating the role of intrinsic vs. extrinsic influences, the influence of how pleasurable past learning experiences have been, the feeling of contributing to something that matters and the importance of having an audience that cares, among other factors (Vroom, 1964; Keller, 1983; Keller, 1987; Ames, 1992; Vail, 1994). Related concepts such as self-efficacy also play a critical role (Bandura, 1977a,b; Pajares 1996; Schunk, 1989; Zimmerman, 2000), and student's beliefs about their efficacy, in turn, influence them emotionally (Bandura, 1997). Several researchers have integrated both affective and cognitive components of goal directed behavior into motivation theories (e.g., Maehr, 1984; Dweck, 1986; Ames and Archer, 1988; Dweck and Leggett, 1988; Elliott and Dweck, 1988). These and many other efforts have provided vast insight into human affect; however, in very few cases are the theories at a level suitable for implementation in an interactive machine model.

The need for more precise theory is being driven today by growing efforts to build technologies that interact with learners—motivating, engaging, and assisting them in challenging new ways. In many of these efforts, the systems need programmed representations and strategies that will perform in real-time interaction with a human learner. The designers of these systems turn to human-human interaction, and its literature, as an example to guide their design. Thus, the intelligent tutoring system research community examines successful human tutoring as a source of inspiration for what might be implemented in machine tutoring systems, and finds, for example, 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" (Lepper and Chabay, 1988). But what do these expert teachers 'see' and how do they decide upon a course of action? The theories, where they do exist, tend to focus on a high-level set of observations and practice, which does not directly translate into the level of detail needed to implement these phenomena into machines.

We propose to test and evolve theories of affect in learning. Our approach will extend classical armchair observations and thought experiments with the development and use of new technologies that help elicit, sense, measure, communicate, and respond to emotions in learning situations. Conducting controlled experiments dealing with affect has always been a challenge, and new technologies are needed to make this process easier. The sections below outline several directions we propose with respect to creating such technologies. Although we treat the technology descriptions separately from the theory in this proposal, we emphasize that the technology development both derives from and contributes to that of the theory; indeed, we hope by engaging in both simultaneously, that one will strengthen the other, bringing both closer to clarity and unity.

Enabling Technologies that Sense and Respond

One of the reasons understanding about affect has lagged behind that of cognition is that affective state information is hard to measure. You can easily measure someone's ability to recall a list of learned items, and with somewhat more difficulty, you can test their ability to generalize and apply some learned information. However, it is much harder to measure how they feel while doing these things. How can various tools of learning, and future robots and environments, sense if a learner is pleased, engaged, disengaged, frustrated or ready to quit? And in what ways can these tools enable reflection and discovery about affect? We propose to develop sensors and interfaces, together with new signal processing, pattern recognition, and reasoning algorithms for assessing and responding to the affect of the learner in real-time.

CHALLENGE: Sensing without interfering

Affective experience—such as how much pleasure, frustration, or interest you felt—is typically measured by questionnaire (e.g., Matsubara and Nagamashi 1996; de Vincente and Pain, 1998; Whitelock and Scanlon, 1996). Special instruments have been developed in many cases, such as for evaluating the motivational characteristics of an instructor's classroom delivery (e.g., Keller and Keller, 1989). Despite the convenience and widespread acceptance of questionnaires, the use of self-report information is considered unreliable when it comes to emotion: For adults, self-report is colored by awareness of internal state, reflections on how such a report will be perceived, ability to articulate what one feels, and more. For children, emotion self-report "is never highly valid, and any report before age 11 is unwise," (Kagan, 2002). On top of these problems, affective mechanisms also include mental functions that people are not typically conscious of, such as intuitions. Finally, the process of administering a questionnaire can be intrusive to the learning experience, changing the very affect that the experience was designed to engender.

A more arduous method of assessing affect is to use external coders to rate a learner's state; this method adds some rigor and reliability, but coding is extremely tedious and subject to its own set of errors. For example, a recent set of experiments that had teachers code three levels of interest plus boredom for children interacting with an educational computer game (Mota, 2002), found agreement among coders to have Cohen's Kappa values averaging 0.79, which is considered "excellent" for that statistic (Bakeman and Gottman 1986). Working with human coders also helps refine the description of observable behaviors associated with each affective state, which is a useful advance as well. However, the use of human coders is almost impossible to scale up to long-term continuous observation of learners. Automating the process would not only free up human coders for more interesting tasks, but would also enable you to assess affect while a learner is learning—which, in turn, enables the technology to respond while there's a chance of affecting the learning outcome. For example, a computer that detects the learner making a mistake while appearing curious and engaged should probably leave the learner alone since mistakes can be important for facilitating learning and exploration; however, if the learner is frowning, fidgeting, and looking around while making the same mistake, then the computer might use this affective feedback to encourage a different strategy.

A number of researchers have raised the concern that you can't begin to measure or respond to affect until after you articulate a clear theory of affect. While theory and clearer terminology are important goals of this Center, there are examples from natural systems that suggest we can still forge ahead, despite the state of the theory. For example, dogs presumably have no theory of what affect is and yet they appear to sense and respond to their owner's moods, responses that in many cases bring about beneficial consequences. One can make a related argument for infants, who show an ability to respond to *how* something is said, long before they understand *what* is said. Thus, we believe that even without a full-fledged theory of affect, machines can be given some capabilities to recognize and respond to affect. In fact, it is our experience that efforts to build a phenomenon that is poorly understood will aid in helping improve the understanding of that very phenomenon, so that engaging simultaneously in both the practice and the theory helps both to make progress.

Emotion recognition is a component of emotional intelligence (Salovey and Mayer, 1990; Goleman, 1995), and skilled humans can assess emotional signals, in themselves and in others, with varying degrees of accuracy. Recent developments in Affective Computing aim to also give computers skills of emotional intelligence, including the ability to recognize emotion as well as a person might (Picard, 1997). The basic approach is to observe a person's patterns of behavior via sensors such as cameras, microphones, or pressure sensors applied to objects the learner is in contact with (mouse, chair, keyboard, steering wheel, toy), and use computers to associate these patterns with probable affective state information. Thus, a camera and computer, equipped with pattern recognition software, might be used to recognize facial muscle movements associated with a smile, and the smile-detection might then be used to help reason about the probability the person is actually happy. (Expressions do not always imply the existence of underlying feelings.) The job of the computer is to assess a constellation of such patterns and relate them to the user's affective state. The latter is what is termed "emotion recognition" even though it doesn't really see what you are feeling, but only a pattern of measurable external changes associated with such.

Most prior work on emotion 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 for recognizing emotions during learning. In general, people can recognize one of about six different emotional states from speech with about 60% accuracy (Scherer, 1981). 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 expressions on groups of 8–32 people (Yacoob and Davis, 1996; Essa, 1997). More recent 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 make up 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 (e.g., Cohn et al., 1999; Bartlett et al., 1999; Donato et al., 1999; Tian et al., 2001; Kapoor et al., 2003; Bartlett et al., 2003) 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. 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 numerical results given 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. Moreover, we are interested in computer recognition of truly experienced emotions in learning situations, as opposed to emotions that have been expressed by actors or by subjects posed in front of a camera or microphone. Thus we cannot expect the computer to perform perfectly at recognition, and our methods will have to take into account uncertainty factors.

Recently, a number of projects have tackled the sensing and modeling of emotion in learning and educational gaming environments (e.g., Conati, 2002; Zhou and Conati, 2002; Zhou and Conati, 2003; Sheldon-Biddle et al., 2003; also see references in the "Prior NSF Awards" section of this proposal.) 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); these are also designed to hide the wearer's expression from view of the teacher, allowing the students to communicate expressions of confusion to the teacher in real time without fear of what the teacher might think. Additionally, the Media Lab has attained from 89–96% classification accuracy of three levels of cognitive-emotional stress (Healey and Picard, 2000), although the latter data were from drivers in Boston and not from children in learning situations, where stress would also be interesting to study.

Recently, a system was designed at the Media Lab for automated recognition of a child's interest level in natural learning situations, using a combination of information from chair pressure patterns sensed using Tekscan pressure arrays (recording how postures moved during learning) (Tekscan, 1997) and from upper facial features sensed using an IBM BlueEyes video camera (http://www.almaden.ibm.com/cs/blueeyes) (Kapoor et al., 2001). New algorithms were developed with the aim of seeing if the machine could match teacher's ratings of affective labels. Training and testing on separate sequences of data from the learning experiences, we developed algorithms that achieve an accuracy of 76% on affect category recognition from chair pressure patterns, and 88% on nine “basic” postures that were identified as making up the affective behaviors. Both sets of results are conservative, being trained on a small set of data, and tested on children not seen before. The accuracy rates increase to 82% and 98%, respectively, on children who have had portions of their data included as part of the training process. All of these results are highly significant, confirming that there is strong evidence of affective information in the postural moves of the child. This same data was used to develop a facial expression recognition system that could operate without any intervention from the experimenter, and without any special calibration steps with the learner. Its performance, training and testing on separate parts of the natural data, was 68% at recognizing six upper facial action units (note that a person has to score 75% to qualify as a human expert). This latter result is conservative, having been attained on children the system hasn't seen before, who were actively fidgeting during the task. These results show that elements of affect can be measured with results significantly higher than random. We propose to further improve these methods, to extend them to recognize lower facial features for smiles and other mouth fidgets, as well as to integrate facial, postural, and other behavioral information for jointly analyzing the state of the learner and increasing accuracy of the inference.

Figure 1: Various sensors can capture postural, facial, skin-surface, and gestural changes that carry affective information. From left to right: Chair with Tekscan pressure sensors, BlueEyes camera, Galvactivator skin conductivity sensor, and Pressure Mouse.

Additionally, we propose to continue to evolve new sensors, for new learning environments where the learner is not seated in front of a computer (see proposed work below) as well as for the traditional keyboard-monitor-mouse environment. The proposed efforts include extending the Galvactivator (http://www.media.mit.edu/galvactivator), a skin-conductivity sensing glove, to communicate wirelessly with a nearby handheld computer. Skin conductivity gives a measure of psychological arousal, which is a strong predictor for both attention and memory (Reeves and Nass, 1996). We have found students enjoyed learning about how this signal changed with their level of engagement in various learning activities (Picard and Scheirer, 2001), and we believe it will be an enabling tool for exploring affect in new learning environments. Additionally, we propose to continue development of a new “pressure mouse” device, a mouse augmented with eight pressure pads that indicate “how” the mouse is being handled. An increase in physical pressure applied to a pressure-sensitive mouse has recently been shown to be associated with frustration caused by poor usability in a computer interface (Dennerlein et al., 2003). We also have additional efforts (via our collaborators at MediaLabEurope) developing new comfortable wireless physiological sensors and real-time signal processing algorithms, which will be useful for monitoring learner stress.

Finally, in neuroscience itself, there is a pressing need to develop ways to measure the levels of interest and motivation of an animal engaged in tasks, and to gauge the levels of stress experienced by the animal. We propose to collaborate with the Graybiel laboratory at MIT to develop new affect sensing technologies for fundamental research looking at the mechanisms of motivation and attention in the animal brain. This research, while very basic, should help inform the understanding of human brain disorders related to attention and motivation.

CHALLENGE: Reflecting and interpreting

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 (1995)

How can systems that measure affective information help people make sense of what has been measured, and respond in useful, appropriate, and respectful ways? One important response is to help people become more aware of their affect—building a kind of an "affective mirror" in which the learner is encouraged to reflect upon how their state is influencing their learning experience. Emotional awareness, in oneself and in others, is considered to be a learnable skill of emotional intelligence. Being aware of one's state, such as frustration, can be instrumental in helping better deal with that state productively.

The Galvactivator, which converts level of skin conductivity to the brightness of a glowing LED, is one device that makes it easy to visualize how your psychological arousal changes as you go about activities. We observed classrooms of students wearing these, where the light glowed brightly when they were engaged in discussing ideas or writing in their journals, and went dim (for many of them) when they were lectured to. Skin conductivity exhibits changes with respect to attention and engagement, and reflects interesting patterns when there are disorders of these, such as in autistics and in patients with various emotion-related deficits (e.g., Bechara et al., 1997; Hirstine et al., 2001). We propose to develop additional means for people to learn about and reflect upon affective signals such as this one.

Figure 2: Prototype of physically animated computer

Another project we propose to explore is the use of a physically animated computer for facilitating learning. This project equips a desktop monitor with the ability to move in subtly expressive ways in response to its user. The physical animation of the machine is inspired by natural human-human interaction: when people work together, they move in reciprocal ways, such as shifting posture at conversational boundaries and leaning forward when interested. The proposed physically animated computer will sense and interpret multi-modal cues from the user via sensors such as those above. It will then respond to the user's cues with carefully crafted subtle mechanical movements and occasional auditory feedback, using principles derived from natural human-human interaction. The initial version of this device will be designed to mirror affect from the user in a way that is non-distracting. For example, if the child's face and posture show signs of intense interest in what is on the screen, the computer would hold very still so as to not distract the child. If the child shifts her posture and moves in such a way that suggests she is taking a break, or starting to become bored, the computer will do similarly. In doing so, the system not only acknowledges the presence of the child and shows respect for her level of attentiveness, but also shows subtle expressions of mirroring that, in human-human interaction, are believed to help build rapport and liking (LaFrance, 1982). By increasing likeability, we hope to facilitate task outcome such as how long the child perseveres with the learning task. Of course, the system can also reflect other expressions such as frustration or disappointment with subtle movements (Liu and Picard, 2003). We are interested in evaluating the impact of such communication on the learner's reflection of her own state, as well as on other performance characteristics of the learning experience.

We are also very interested in integrating new affect sensing, recognition, and reflection technologies into efforts to build intelligent tutoring systems and other automated systems where there is potential to adapt the learning experience based on signs of interest, frustration, and any other affect-related cues. We currently have one such collaboration with the AutoTutor project at Memphis (http://www.autotutor.org), and we propose to make available resources to facilitate other researchers' efforts to integrate new affect sensing technologies into such efforts.

By embedding these technologies in learning interactions with automated systems (animated tutors, robotic computers, etc) and also integrating them into other learning environments (see below) we hope to better answer such questions as: What affective states are most important to learning and how do these states change with various kinds of pedagogy? How does knowledge of one's affective state influence outcomes in the learning experience? Additionally, these technologies form the basis for building systems that will interact with learners in more natural ways, bootstrapping the machine's own ability to learn, the topic of the next section.

OUTCOMES

Years 1–2: Complete design, test, and debugging of wireless skin conductivity sensor and new mouse prototypes. Improve facial expression recognizer to include lower-face action units and features that are more invariant to head motion. Integrate pressure-sensitive mice, facial expressions, posture, and behaviors of learner into inference algorithms. Improve pattern recognition techniques for integrating multi-modal noisy data and for learning how to trade-off features over time. Work with co-learning projects (below) and others crafting new learning environments, to help design and develop sensors for their needs. Finalize robotic computer platform; refine and user-test its mirroring movements for distraction and comprehension.

Years 3–5: Similar to Years 1–2, but sensors, algorithms, and responses will be specialized and adapted to support the projects described below. In some cases, sensors will need to be redesigned for directly integrating into new haptic tools and learning environments, e.g. instead of a pressure-sensitive mouse, the haptic input devices can be made pressure sensitive. For other applications we will design analysis tools for EEG and ECG to improve the ability to monitor attention and stress under natural learning conditions. The design of affect-mirroring capabilities will be integrated into the learning companion system (described below.) Throughout these efforts, our goal will be to construct useful, accurate, re-usable, and robust tools for sensing and responding to affect.

Modeling Affect and Cognition in Co-Learning Systems

By modeling affect and cognition in systems that learn with people through natural interaction, we hope to realize three equally important goals. First, we wish to advance the state of the art in machine learning to develop systems that can learn far more quickly, more broadly, and continuously from natural human instruction and interaction than they could alone. Second, we aspire to achieve a deeper understanding of human learning and development by creating integrated models that permit an in-depth investigation into the social, emotional, behavioral, and cognitive factors that play an important role in human learning. Third, we shall use these models and insights to create engaging technologies that help people learn better.

The history of combining affective mechanisms with cognitive ones for improving machine processing goes back at least to the work of Herb Simon (1967), who articulated the construction of motivational and emotional controls over cognition, and proposed incorporating these into information processing systems. His work was inspired by Neisser (1963), who in a criticism of the dominant information processing theories, emphasized these fundamental characteristics of human thought: (1) human thinking always takes place in, and contributes to, a cumulative process of growth and development; (2) human thinking begins in an intimate association with emotions and feelings which is never entirely lost; and (3) almost all human activity, including thinking, serves not one but a multiplicity of motives at the same time. Some have claimed that the emotional components may be, in some ways, dominant: Don Norman wrote, "There must be a regulatory system that interacts with the cognitive component. And it may well be that it is the cognitive component that is subservient, evolved primarily for the benefit of the regulatory system, working through the emotions, through affect" (Norman, 1981). Thus, these pioneering theorists suggested that cognition, and by extension learning, took place within the context of emotional, motivational, perceptual and behavioral structures that shaped those very processes.

By contrast, the dominant trend in machine learning has been to eschew built-in structure or a priori knowledge of the environment and to discover structure that is in the data or the world through exhaustive search and/or sophisticated statistical learning techniques. Pattern recognition and reinforcement learning are two problem domains in particular that have attracted attention, and met with good success. In pattern recognition, the system's goal is to learn a mapping from a set of input features to an output "label." The input features might be associated with a gesture, a face, or an acoustic pattern, for example, and the label might be a something like "appears to be happy." The system typically learns the mapping through a statistical analysis of hundreds or thousands of training examples chosen by a "knowledgeable external supervisor" (Sutton et al., 1998), in which an example contains both the input features and the desired output label. Typically, the system has no a priori knowledge of the structure of the input space and must discover it based on the examples provided by the supervisor. In the domain of reinforcement learning, the goal of the system is to learn an optimal sequence of actions that will move the system from an arbitrary state to a goal state. The main approach of reinforcement learning is to probabilistically explore states, actions and their outcomes to learn how to act in any given situation.

Reinforcement learning is an example of unsupervised learning in that the only supervisory signal is the reward received when it achieves the desired goal. However, as with supervised learning techniques, the actual learning algorithm has no a priori knowledge about the structure of the state and action spaces and must discover any structure that exists on its own through its exhaustive exploration of these spaces. As a result, reinforcement learning typically requires hundreds or thousands of examples, in order to learn successfully.

Thus, the progress to date in machine learning has come with some caveats. First, the most powerful techniques rely on the availability of a great deal of data. Thus, they are often not appropriate in domains in which the number of examples is very small. Second, they tend not to be appropriate when the environment is changing so quickly that earlier examples are no longer relevant. Third, the underlying representations used in machine learning typically make it difficult for the systems to generalize from learning one particular thing or strategy to another type of thing. Fourth, little attention has been paid to the question of how a human can guide the learning process. Fifth, and not insignificantly, few would argue that current approaches to machine learning, however successful, have much to tell us about how learning occurs in animals and humans,

By contrast, any survey of animal learning will quickly convince one that learning in nature is characterized by fast and robust, albeit, constrained learning (Shuttleworth, 1998; Gould and Gould, 1999). For example, a dog can be trained to roll over in response to an arbitrary verbal or gestural cue in as little as 20 to 50 repetitions (Wilkes, 1995). A Nightingale can learn to imitate the song of another bird after as few as five presentations (Marler, 1990). A typical child learns an average of 8–10 words a day over their first five years (Markham, 1990). How is it that animals and children can solve these learning problems so effortlessly?

Our hypothesis is that the answer lies not in finding the “magic bullet” of a unitary learning algorithm or module, but rather in discovering the combination of underlying structures and processes that radically simplify what would otherwise be a complex learning problem. In nature, these internal structures are cognitive, behavioral, social, emotional, motivational, shallow and deep, innate and learned, purposed and repurposed. Indeed, an important way that internal structures simplify the learning task is by acting so as to bias the learner to take maximal advantage of external environmental and social-emotional interactions that serve to structure and constrain the learning task. Hence, learning is the result of a complex interplay of structures and processes, both internal and external to the learner, and having both cognitive and affective aspects.

We propose to develop computational models and learning systems that capture these key characteristics. We believe that such models can provide new insights into numerous cognitive-affective mechanisms, and shape the design of learning tools and environments, even if they do not compare to the marvelous nature of those that make children tick. The three co-learning scenarios presented below illustrate the kinds of systems we propose to build: Each one learns in partnership with people, but with different emphases. Following these, we present core challenges in cognitive-affective co-learning.

A Curious Robot

Imagine a robot that exhibits curiosity (Figure 3). Curiosity is a trait of natural learning systems (i.e., people and animals) that exhibit inquisitiveness and a drive to learn, which tends to be followed by quickly learning what they ought to learn, when they ought to learn it, in an ongoing way. Inspired by nature, we envision a Curious Robot to be a pro-active learner that seeks out experiences and people from which to learn new things.

Humans are natural and motivated teachers for entities that are rewarding to work with. A Curious Robot will be able to leverage the rich social nature that is uniquely characteristic of human learning and instruction to constrain and bias its own exploration and discovery of new skills and knowledge—thereby allowing itself to learn more quickly, broadly, and continuously than it could alone. To do so, a Curious Robot will need a deeper understanding of the learning process, beyond turning the statistical crank of a learning algorithm, to actually reflect upon the learning process: when to learn (or when to get help to learn), what to learn, from whom, how (e.g., recognize success, correct errors, judge progress), and why? In humans and animals, both cognitive and affective factors play an important role in this process. By building this robot, we would further illuminate these factors.

We will consider such a robot to be successful if it can learn a variety of new skills, tasks, and knowledge from natural human instruction, without requiring any adjustment of the internal learning mechanisms. We believe this will go a long way to enabling a new class of technological artifacts that readily adapt and learn within the human environment.

A Teachable Interactive Character

Next, imagine a scenario in which a child teaches a 3-D computer-animated puppy new tricks. Teachable agents are a new area of research (Biswas et al., 2001a,b) and show promise not only in motivating learners, but also in engaging them in opportunities to reflect on attitudes about learning and other meta-learning concerns. Animal training can be viewed as a coupled system in which the trainer and the animal cooperate so as to guide the animal's exploration to discover how to perform new skills. Animal trainers have developed techniques such as "luring," "shaping" and "clicker training" that allow the person to guide the animal's learning from its observed behavior alone (Pryor, 1999; Lindsay, 2000). Because the trainer cannot see inside the animal, it is very important that the virtual puppy's behavior be an immediate and accurate reflection of what it has learned so far. Moreover, to the extent it can infer, even minimally, the trainer's intent, and use that knowledge to guide what it learns from the trainer, it will be markedly easier to train, and learn in far fewer examples. This will provide the child with immediate and compelling feedback as to the success or failure of his or her teaching efforts.

This engaging experience will give the child an opportunity to learn about the importance of exploration, motivation, context, timing, and built-in biases for efficient learning. By putting the child in the position of helping to guide the puppy's learning, we hope to encourage the child to ask questions about his or her own learning and consider how human learning and teaching may be similar but also different from animal training in interesting ways. Such a system will have succeeded if the child gains a deeper understanding of the teaching-learning process and uses it to adapt his or her behavior and hypotheses about learning and teaching.

A Learning Companion

Finally imagine a scenario in which a machine serves as a computerized Learning Companion to facilitate a child's own efforts at learning. A Learning Companion will not be an intelligent tutoring system that already knows the answers about the subject being learned, but rather it is a player on the side of the student—a collaborator of sorts—there to help the child learn, and in so doing, learn how to learn better. To do so, the companion will help to 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. 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.

Building the kinds of co-learning systems envisioned above presents a number of interesting and challenging problems. Below, we outline our approach for how to address several of the challenges involved in building co-learning systems that can learn quickly, broadly, and continuously in partnership with a person. This work will not only contribute to building more effective co-learning systems, but also allow us to thoroughly explore the role of cognitive, affective, social, and behavioral factors in the co-learning process.

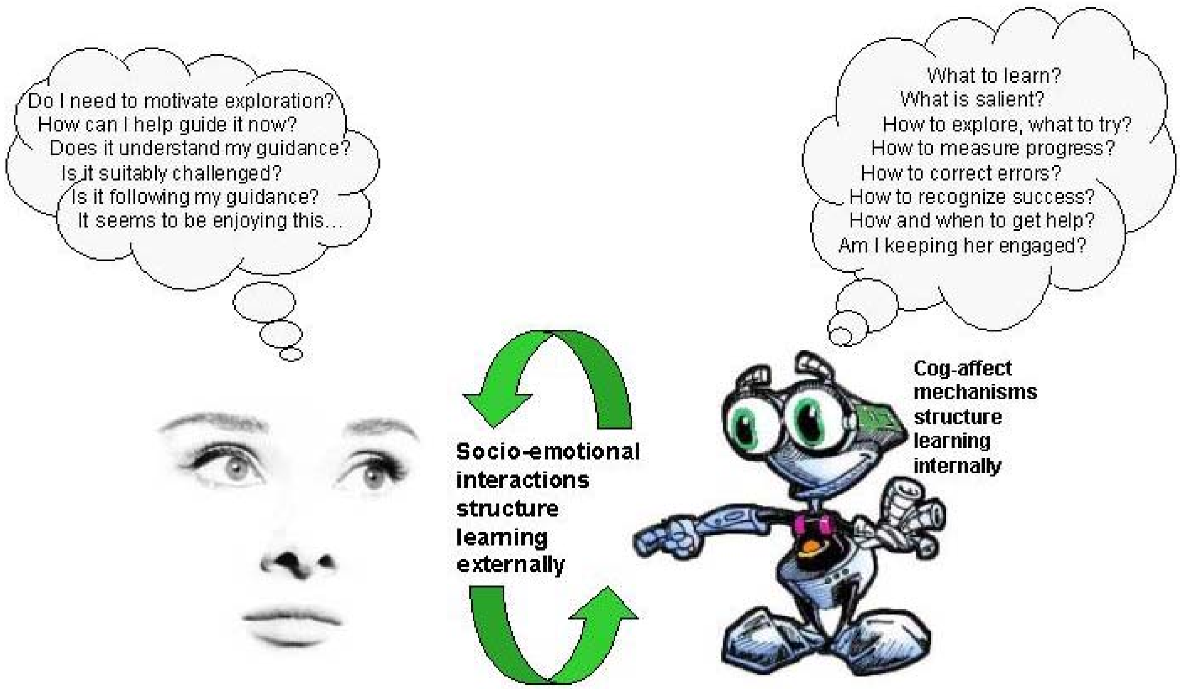

Figure 3: A co-learning system involving cognitive-affective mechanisms that regulate both internal and external processes

CHALLENGE: Learning quickly from few examples

We propose to model internal cognitive-affective structures that work in concert with perceptual and affective information gained from real world interaction, to dramatically accelerate a system's ability to learn the right thing from a limited number of examples. We have demonstrated this approach in domains as diverse as learning word boundaries from fluent speech (Roy, 2000; Roy and Pentland, 2002), teaching animated characters new tricks (Blumberg and Downie, 2002; Blumberg 2002), and learning how to imitate facial expressions (Breazeal et al., 2003).

For instance, inspired by a close study of the practice of animal training (Pryor, 1999; Wilkes 1995), we use behavioral context and internal biases to dramatically speed up learning from few examples. Consider the problem of associating a verbal cue with a behavior. This requires the learner to build a perceptual model of the cue based on examples provided by the trainer in the course of training. How does the system know when an example has been given and whether it is a good example of the cue or not? Our animated dog, Dobie, uses rewarded actions as the context for identifying important sensory cues and for guiding the construction of a perceptual model of the cue. For instance, a good example of the acoustic pattern "sit" is one that occurs just before or during a sit action that results in reward. By using natural feedback signals to guide the construction of perceptual models, and by only attending to stimuli that occur within a narrow temporal window, fewer models are built, and those that are built tend to be more reliable and robust. These, and other biasing mechanisms, enable learning to happen with one to two orders of magnitude fewer training examples (Blumberg and Downie, 2002; Blumberg 2002.)

We provide trainers with a "clicker" that allows them to precisely mark the behavior they wish to encourage or discourage. In addition, Dobie's learning mechanism incorporates heuristics that attempt to match the trainer's expectation as to what Dobie "should be" learning from the trainer's actions. For example, if the trainer manipulates Dobie into a "down" posture and then rewards him, the down action gets the credit for the reward, rather than the action that was active while the trainer was luring him into the posture, e.g., the "follow-food-in-hand" action. Similarly, Dobie always assigns reward that he receives to the behavior that is currently observable to the trainer. Finally, Dobie assigns credit to the most novel and relevant combination of perceptual cue and action. For example, suppose Dobie decides to sit, and as he begins to do so, the trainer says "sit" and subsequently rewards him. Even though Dobie chose to sit spontaneously, he will assign the reward to sitting in response to being told to do so because that is a more novel combination. These are just a few examples of structures that dramatically speed up learning.

We propose to develop these ideas further in our Curious Robot project and apply them to social interaction cues that people use to teach each other. For instance, the "click" in clicker training is used to signify both salience as well as to signal positive affective value. Extending this work, we shall model cognitive-affective mechanisms of attention (Breazeal et al., 2001) and saliency measures to allow the robot to determine the significance of stimuli or events, either on its own or when guided by a person (via gesture or directed gaze). When learning something new, it is also important for the robot to assign affective value to incoming stimuli to help bias what it learns, e.g., Was the outcome good or bad? Am I making progress toward a desired outcome? Does my co-learner appear pleased? etc. We shall model internal cognitive-affective mechanisms to assess the affective value of both internal and external states. For instance, social referencing (where an infant looks to the expressive reaction of his or her caregiver) plays an important role in helping an infant (as it could for a robot) to affectively evaluate novel situations and to guide his/her subsequent exploration (Feinman et al., 1992). This will allow our system to learn quickly from natural social interaction, and allow us to explore cognitive/affective models of saliency and affective value with respect to learning in people and animals.

CHALLENGE: Learning broadly with human guidance

We propose to give a robot the skills necessary to take advantage of natural human instruction to guide its own exploration and discovery (Breazeal, 2002). Humans are often natural and gifted teachers, capable of providing the appropriate amount of guidance to help foster learning the right thing at the right time. Much of human instruction and teaching consists of subtle and not-so-subtle communication and cues that help to constrain and facilitate the learning task. For instance, *scaffolding* as traditionally viewed by the field of developmental psychology, is conceptualized as supportive structures provided by an adult whereby the adult manipulates the child's interactions with the environment to foster novel abilities (Wood et al., 1976). Parts of scaffolding, such as reducing the number of degrees of freedom in the target task, and either demonstrating or guiding the learner to perform particular action, may be considered cognitive structuring, while other parts such as reducing distractions, drawing attention to the task's critical attributes, and giving the child positive/negative forms of feedback and encouragement, involve affective structuring.

In previous work, we have explored this process through modeling the development of facial imitation with an expressive robot, Leonardo (Breazeal et al., 2003). An increasing amount of evidence suggests that the ability of humans to learn by watching others, and in particular, the ability to imitate, is not only important for learning new manipulation and problem solving skills (Byrne and Russon, 1998), but it could also be a crucial precursor to the development of appropriate social behavior, and ultimately the ability to reason about the thoughts, intents, beliefs, and desires of others (Meltzoff and Gopnik, 1993, Meltzoff and Moore, 1997). Bi-directional imitative games allow the infant to share in the anticipation of some simple and predictive sequences of events that are under the infant's own voluntary control given a cooperative adult. This in turn gives the child even finer control over adults' behavior, so that the infant can explore and gain further information and learn more models of motor and communication skills. Inspired by this process, Leonardo learns how to imitate the facial expressions of people during playful imitative exchanges that capture the bi-directional nature of early infant-caregiver imitative games. We propose to extend this work to scaffold the robot's ability to learn both social and object-manipulation skills, and to apply these skills to learn multiple tasks from natural interaction with a person.

As discussed earlier, affect plays an important role in influencing how "broadly" or "narrowly" people learn, and we propose to explore this dynamic through computational modeling. In addition, the robot will need the ability to learn in the context of servicing multiple goals—trying to imitate you, trying to apply what it has learned to a new situation, trying to get you to play with it, etc. It is argued that affective and cognitive mechanisms play an important role on regulating and prioritizing goals in animals and people (Norman, 1981), as they will for our robot. From this body of work, we will not only advance the state of building machines that can learn from people, but also investigate relationship that social, cognitive, and affective factors play in this broad style of learning.

CHALLENGE: Diverse learning over the long term

We propose to develop a computational model, inspired by behavioral development in animals, that seeks to manage, perform and integrate diverse "learning" strategies within a long-term behavioral context.

Animal learning throughout development is a highly discrete process, consisting of overlapping episodes of highly specialized learning (Coppinger et al., 1990; Lorenz et al., 1973; Gottlieb, 1991; Hogan, 1999). Indeed, animals are adept at learning when the environment is ready to teach, as if they were "looking for" the requisite information when it is most likely to be available and easiest to acquire. In addition, important behaviors are often self-motivating, and are later repurposed for their ultimate use. Thus, a kitten perfects its pounce, not on prey, but on its littermates, apparently because they are enjoyable to pounce on. Later, the motivational context dictates the object and form of the pounce: for pleasure, the pounce may aim for an object that is particularly "satisfying" on which to pounce (and perhaps repeatedly so); when hungry, the cat chooses the object and form of pounce that maximizes the likelihood of satisfying hunger (Lorenz et al., 1973).

This work will focus on the concept of "specialized learning structures." These structures will include mechanisms that identify when to begin learning, the learning strategies to be employed, components of motivation so as to engage in the behavior required to facilitate the learning, fail-safe mechanisms in case the appropriate learning conditions don't arise, and ultimately a mechanism that identifies when to stop learning. These structures will be embedded in an overall behavioral context that undergoes re-organization as a result of the new skills being integrated in, or being repurposed for different motivational goals.

Another aspect of continuous learning focuses on sustaining the long-term and reoccurring interaction of the learning system with other co-learners. Much of human instruction and scaffolding is not only helpful in simplifying the learning problem, but is an enjoyable experience for both the human teacher and learner. In our experience, however, most computational agents in existence are at best charming for a short interaction, but rapidly grow tiring or annoying with longer-term use. In order to build an agent that people will enjoy over time, it must be able to both accumulate memory of ongoing interactions with the user, and also exhibit basic social-emotional skills. We propose to address these issues in all three of our proposed projects.

Going a bit further, it is known that the presence of someone who cares, or at least appears to care, can be motivating (Wentzel, 1997). Various studies have linked interpersonal relationships between teachers and students to motivational outcomes over the long term (Pianta, 1992, Wentzel and Asher, 1995, Birch and Ladd 1996). What is it that students think teachers do to give the impression of caring? One suggestion is that the teachers (1) model caring behavior; (2) expect as well as encourage students to do the best they can given their abilities; and (3) engage students in dialogues that lead to mutual understanding and perspective taking (Noddings, 1992). Although computers don't "care" in the sense of having feelings like people have, it is nonetheless possible for them to model certain behaviors and give some of the other impressions that contribute to a perception of caring. For example, Bickmore designed, built, and tested a "relational" agent that had certain social-emotional skills for crafting long-term relationships with people, and found that this relational agent (compared to an identical agent without relational skills) led to statistically significant effects where people reported that the relational agent cared more about them, was more likeable, showed more respect, and earned more of their trust than the non-relational agent. People interacting with the relational agent were also significantly more likely to want to continue interacting with that agent (Bickmore, 2003; Bickmore and Picard, 2003, also described in the "Prior NSF Awards" section below.)

Our relational agent experiment was not explicitly about caring; nonetheless, it demonstrated that it is now possible to begin to functionally tease apart behavioral elements of notions such as caring, and to build interactions where these elements are systematically included or excluded. Such interactions would be of interest in examining the effects of various kinds of caring presences on learning—and do so in a controlled way that is virtually impossible to do with a real person. Although we do not expect that machine "caring" could provide any kind of real substitute for genuine human caring, we do hypothesize that certain aspects of it could be given to learning tools and technologies in ways that positively impact learners. We propose to investigate this issue as part of the evaluation of our Learning Companion project.

OUTCOMES

Our primary deliverable will be a cognitive-affective architecture with the capabilities described above, and a number of instantiated systems that showcase the architecture's capabilities (e.g., a Curious Robot, a Teachable Interactive Character, a Learning Companion). The impact of these systems will be evaluated through carefully designed user studies and findings will be submitted for peer-reviewed publications.

Year 1: (1) Design the specifications for the cognitive-affective architecture including learning models that incorporate cognitive, affective and social factors. (2) Integrate *existing* technologies for real-time perception of human social cues, speech and gesture, and motor control and computer animation into the overall architecture. (3) Fabricate necessary hardware systems and start development of new perceptual abilities needed for the Curious Robot, Learning Companion, and Teachable Interactive Character projects. (Portions of this latter goal will collaborate with the sensing projects above.)

Year 2: (1) Refine sensing and perceptual systems for the three proposed projects, demonstrate their performance, and integrate them into the cognitive-affective architecture. (2) Integrate models of emotion, motivation, and other affective and cognitive abilities into the cognitive-affective architecture. (3) Implement and demonstrate initial learning and teaching test scenarios for each of the three proposed systems. (4) Design an experiment to evaluate the performance of the Teachable Interactive Character.

Year 3: (1) Refine the design of the cognitive-affective architecture and its mechanisms. (2) Demonstrate the ability of a Curious Robot to learn a single task faster from natural human instruction than it could alone. (3) Conduct an experiment to evaluate the impact of the Teachable Interactive Character on what children learn about learning. (4) Design an experiment to evaluate the effectiveness of the Learning Companion on emotion reflection and attitudes.

Year 4: (1) Continue to elaborate on the design of the cognitive-affective architecture and its mechanisms. (2) Demonstrate the ability of a Curious Robot to learn multiple tasks from natural human instruction from one continuous interaction. (3) Conduct an experiment to evaluate the impact of the Learning Companion on traditional learning outcome measures.

Year 5: Conduct and experiment to evaluate the learning performance of the Curious Robot (quicker, broader, more continuously) with subjects who are simply asked to teach the robot how to perform 2 to 3 sufficiently different tasks with no prior coaching of the subject by the experimenter. Examine the effects of various components of "caring behaviors" on long-term learning interactions.

Fostering Deep Engagement

An interest is a terrible thing to waste. —Roger Schank (1994)

Interests are a great untapped resource. In schools of education, many courses emphasize how and what teachers should teach, but rarely examine why their students might want to learn. When the issue of motivation is addressed, the emphasis is often on extrinsic motivators and incentives, such as grades and prizes based on performance. Yet if you look outside of school, you can find many examples of people learning—in fact, learning exceptionally well—without explicit "rewards." Youth who seem to have short attention spans in school often display great concentration on projects that they are truly interested in. Research has found that "self-motivation, rather than external motivation, is at the heart of creativity, responsibility, healthy behavior, and lasting change…The proper question is not 'how can people motivate others?' but rather 'how can people create the conditions within which others will motivate themselves?'" (Deci, 1995).

Researchers have begun to map out the conditions under which people become most deeply motivated and engaged. Csikszentmihályi has highlighted the dynamic between "challenge" and "mastery" in the learning process (e.g., Csikszentmihályi, 1991, 1997). Too often, designers and educators try to make things "easy" or "entertaining" for learners. But Csikszentmihályi has found that people become most deeply engaged in activities that are challenging, but not overwhelming. Similarly, Papert has found that learners become deeply engaged by "hard fun"—that is, learners don't mind activities that are "hard" as long as the activities connect deeply with their interests and passions (Papert, 1993).

Researchers have found that new media technologies have great potential for fostering deep engagement. The issue here goes beyond the obvious appeal of interactivity and rich visual effects. New media technologies make possible new representations—which, if properly designed, can connect more closely and deeply with learners' interests, intuitions, and experiences. The Logo turtle is a classic example. The turtle offers a new framework for learning geometry (based on ideas from differential geometry), in sharp contrast to the traditional Euclidean formulation. With this new representation, children can learn important geometric ideas in a more "body-syntonic" way, imagining themselves as the turtle as it draws out geometric shapes and patterns, and thus leveraging their intuitions and experiences their own bodies (Papert, 1980). This example highlights an important design strategy: to re-purpose the "deep structures" of the mind (in this case, structures based on experience with one's own body) to support the learning of new things.

Through these types of activities, new media technologies have the potential to "make the abstract concrete" (Turkle and Papert, 1990). A growing number of researchers (e.g., Lave and Wenger, 1991) have argued that people form their strongest relationships with knowledge through "concrete" representations and activities—very different from the formal, abstract representations and approaches favored in traditional school curricula (particularly in the domains of math and science). At a minimum, new media technologies open opportunities for multiple representations of knowledge. This diversification of representations is critical to engaging more students in productive math and science learning: students who have styles of thinking and learning that are not well matched to traditional representations often find alternative representations more accessible, engaging, and empowering.

Over the past decade, researchers and educators interested in fostering deep engagement in the learning process have focused increasingly on design-based activities, in which students create external artifacts that they can share and discuss with others (e.g., Lehrer, 1993; Soloway et al., 1994; Kolodner et al., 1998). In particular, Papert (1993) has laid the foundations for a "constructionist" theory of learning and education, extending constructivist theories of Piaget and others by arguing that learners "construct" new knowledge most effectively when they are in the process of "constructing" meaningful artifacts in the world. Constructionist activities engage people as active participants, giving them a greater sense of control over (and personal involvement in) the learning process. Moreover, the artifacts that people design serve as external shadows of the designer's internal mental models. These external creations thus provide an opportunity for people to reflect upon—and then revise and extend—their internal models of the world.

The Center for Affective Learning will build on these ideas, both by (1) developing new theoretical frameworks for understanding the nature of "deep engagement," and (2) developing and studying a new generation of technologies, activities, and environments that are designed explicitly to foster deep engagement with important domains of knowledge and with the learning process itself.

CHALLENGE: Bringing together the physical and the digital

People grow up in a physical world. They develop sophisticated skills and intuitions for sensing and manipulating their physical environments. Early-childhood educators, following the tradition of Froebel and Montessori, have recognized the importance of this physical-world knowledge and body knowledge, using manipulative materials as a means for children to explore and learn important concepts. Some historians argue that Froebel's "gifts" (the manipulative materials developed for the first kindergarten in early 19th century) deeply influenced the course of 20th century art; indeed, Frank Lloyd Wright credited his boyhood experiences with Froebel's gifts as the foundation of his architecture (Brosterman, 1997). But most efforts to use digital technologies in education have ignored the physical world, and thus missed important opportunities to leverage learners' existing knowledge.

The Media Laboratory has been a leader in integrating the physical and the digital, pioneering the field of "tangible interfaces" (Ishii, 1997) and introducing the idea of "digital manipulatives" (Resnick, 1998). Our "programmable bricks" (commercialized as LEGO MindStorms) have been used by millions of students around the world. Even a superficial eye can see that learners are more engaged when they learn principles of physics and engineering by building functioning machines. Our research has shown that this engagement comes, in part, from learners' personal "identification" with the robots and machines that they build (analogous to their identification with the Logo turtle). In our NSF-funded Beyond Black Boxes project, we found that students, by building their own robotic constructions, not only became more motivated in science activities, but also developed critical capacities in evaluating scientific measurements and knowledge, made stronger connections to the scientific concepts underlying their investigations, and developed deeper understandings of the relationship between science and technology (Resnick et al., 2000).

Building on this foundational work, we propose to develop a new generation of technologies that make abstract concepts more manipulable and thus more learnable—not just in early-childhood education (as with traditional manipulatives) but for learners of all ages (including university students, designers, and business professionals). Our goal is to develop new technologies that support representations that are better matched to learners' intuitions, interests, and ways of thinking—enabling learners to explore complex concepts (such as system dynamics) that were previously viewed as too advanced or difficult. We will also facilitate collaborative learning with technologies that support simultaneous access and manipulation of "tangible representations" by groups of learners. At a more fundamental level, we aim to develop a better theoretical framework for "tangible thinking," analogous to previous research on "visual thinking."

CHALLENGE: Bodies of knowledge

With traditional computer interfaces, many users end up sitting hour-after-hour in front of screen and keyboard. Even most mobile technologies are not designed for use on the go: they call for quiet use of fine motor skills while looking at a small display. We propose to counter with new technologies and activities that support and encourage more active use of the body in the learning process. We will explore how expressive use of one's own body, and attention to the nature of one's own movements, can serve as bases for new types of learning experiences.

We will work toward understanding the richness of human movement so it can be incorporated into the design of interactive tools and environments. We will address movement in different contexts and at different spatial and temporal scales—from fine-motor to gross-motor, from involving fingers and hands to whole-body movements, and from simpler to more complex, requiring shorter or longer periods of time to both learn and perform the actions (Gawande, 2002; Turvey, 1989). The movements may themselves support development of particular learning strategies. Knot-tying, piano playing, juggling, skiing, and dance are example domains in which the body in motion can support intuitive learning about apparently unrelated but potentially deeply interconnected conceptual realms (Austin, 1974; Burton et al., 1984; O'Modhrain, 2000; Strohecker 1991, 1999).

We are particularly interested in exploring how dance can serve as a basis for developing mathematical and scientific intuitions. In summer 2003, we organized a RoBallet workshop in which a group of children (ages 9–12) choreographed and performed dances in an interactive space of their own design. The children decided where to place sensors, both on the stage and on their bodies, and then used their movements to control and alter music that they had composed, robotic devices they had built, animations they had created, and lights that they had programmed. We expect there to be many fruitful extensions of this research project. We will explore whether planning and thinking about one's own movements through 3-dimensional space can provide a foundation for thinking like a geometer or topologist, and whether choreographing and coordinating sequences of movements can provide a foundation for understanding formal systems and thinking like a computer programmer.

CHALLENGE: New spaces for creative expression and collaboration

We propose to study how new computing and communications technologies create opportunities for new types of "design spaces" and "learning spaces"—and, in the process, transform people's relationship to (and attitudes about) the learning process. Our goal is to develop a new category of spaces and places that combine activities and attitudes traditionally segmented between schools, museums, videogame arcades, theme parks, performing arts complexes, and community centers.

In this effort, we are inspired by the innovative design of the pre-schools in Reggio Emilia, Italy (Malaguzzi and Ceppi, 1998). We hope to extend these ideas to new audiences and new materials; in particular, we hope to create spaces where new digital materials have the same status as the manipulative and craft materials at the heart of the Reggio classrooms. In developing and studying these new spaces, we will draw on Media Lab expertise and experience in responsive environments, collaborative design, and tools for expression and invention. In particular, we will build on the lessons we have learned from successful projects such as Toy Symphony, the Computer Clubhouse network of after-school learning centers, and the NSF-funded Playful Invention and Exploration (PIE) network of museums.

New technologies provide opportunities to rethink the ways we organize spaces. When we created the first Computer Clubhouse a decade ago, there were no wireless or mobile technologies, so space was organized around desktop computers. Today, very different configurations are possible. This fall, in collaboration with the Aga Khan Foundation, we are organizing a month-long charrette in which graduate students at six leading schools of architecture will develop designs for a “next generation Computer Clubhouse.” A new Computer Clubhouse will be built in Kenya based on the winning submission.

To continue to test and refine our ideas about the design and use of space in technology-rich environments, we will develop an experimental “Garden for Creativity and Collaboration” (GCC) within the new Media Lab building (being constructed adjacent to the current Media Lab facility, scheduled to open in 2006). We will use the GCC for workshops and other public programs, providing opportunities for young people from surrounding communities to play, experiment, and create with new Media Lab technologies—and opportunities for Media Lab researchers to rethink and iteratively redesign our spaces, technologies, and activities.

CHALLENGE: Motivational music

“Music affects *the emotions by* presenting *us with emotions; and it is not surprising that the effect of music on us is often not dissimilar to situations in which we observe strong emotion in other humans. Of course, language can present the emotions too, sometimes with powerful effect. But music does it ever so much more easily, just because…first, it can present ranges and subtleties of feeling that conventional language cannot, and, second, that it is a quasi-natural and not just a conventional representation of feelings.”* —Laird Addis (1999)

We propose to study how music affects human behavior, and how active involvement in music-making—facilitated by new interface and interaction technologies—can exert a powerful influence on motivation and learning. A variety of recent studies have shown that students who learn an instrument perform better in math. It is also known from clinical case studies that music can affect—in very specific ways—human neurological, psychological, and physical functioning in areas such as processing language, expressing emotion, memory, and physiological and motor responses (Tomaino, 2002). However, music still tends to be used in primarily recreational ways, with little understanding as to how it contributes to learning.

Our goal is to develop a comprehensive set of musical tools and techniques to investigate whether (and, if so, under what conditions) musical activities are associated with enhancements in memory, concentration, and imagination. We will focus especially on technologies that enable novices to express themselves musically (in the tradition of our HyperScore software), studying how new expressive abilities are associated with improvements in motivation and learning. In this effort, we will bring together Media Lab expertise in music cognition and perception, music composition and design, hyperinstruments for novices and experts, synthetic performers and listeners, expressive performance sensing, and affective interfaces. In addition, we will launch new initiatives in haptic playback and enhancement, musical neurology, and intelligent audio analysis and reconstruction (with an emphasis on augmented vocalization and the incorporation of everyday sounds into musical contexts). For example, in our haptic playback research, we will record the sophisticated gestures used by expert musicians (in bowing a violin, for instance), then use that information to physically guide a novice musician's body through the same gesture, with special algorithms to intelligently emphasize salient points (providing the learner with an exaggerated, dramatized feel of particular aspects of the gesture). We expect that this system will lead to discoveries, techniques, and theories that can be applied in many other situations, including communications, sports, and focusing exercises.

Building on successful Media Lab projects such as the Brain Opera and Toy Symphony, and our ongoing relationships with arts and educational institutions (including the Boston Symphony Orchestra, Carnegie Hall, the Los Angeles Music Center, Seattle's Experience Music Project, the Institute for Music and Neurology at Beth Abraham, and the New York City Public Schools), we will test our new musical activities in a wide variety of artistic, educational, and community environments. And although we will primarily concentrate on the uses of music in enhancing learning for young people (ages 2–18), we will also explore music's potential for augmenting learning at any age, as well as for contributing to the success of learners with intellectual, emotional, or physical limitations.

CHALLENGE: Roots, fruits and shoots

Learning is rooted in the person and the culture; it bears fruit through the construction process; and it has shoots that branch into new areas. These principles of learning are experientially based, differing markedly from the concept that requires a disconnected accumulation of chunks of knowledge. In order for the learning to become truly rooted, a person has to have a deep emotional attachment to the subject area. Rooting and the possibilities for branching flow from a better understanding of affect, comfort, culture and motivation.

We have found that when individuals participate in shared activity, they not only develop new ideas for themselves *(fruits*), but the collaborative process transforms and shapes the organization of the community as a whole (*shoots)*. So rather than thinking of learning and expertise as a body of acquired skills and knowledge unique to an individual, we adopt a more "situated learning" perspective, where the unit of analysis for expertise is the community itself.

Researchers (e.g., Wentzel, 1997) argue that the pursuit of prosocial and socially responsible goals have a positive impact on student perception of caring in a classroom setting. The efficacy of a virtual expert within the classroom has also been demonstrated (Evard, 1996): students and community members share information more readily and take more risks. But do these results transfer outside of the classroom? Chesnais (1999) argues that skills, support, and effort are necessary for successful community expression and construction. In the absence of human mentoring and an explicit mechanism for external support, can affective-learning technology help people find the latent expertise and share support within and across the communities?

Digital technologies offer new opportunities for discovering roots, adapting to preferences, and enabling creative and idiosyncratic connections to learning and knowing. We are developing a new range of expressive technologies and a more integrated methodology to facilitate rooted knowledge construction and support development of shoots to new areas through electronic collaboration and support (Cavallo, 2003).

An example is found in our work discovering engine culture in Thailand (Cavallo, 2000). Numerous local innovations and widespread knowledge made it clear there was a deeply rooted culture of learning and practice building upon knowledge from the internal combustion engine. This became evident in our Project Lighthouse when rural adolescents, all of whom had left school after only a few years, used a variety of computational technologies to design a new dam and address critical water problems in the region. Not only was it remarkable that they learned enough to design an irrigation system without the usual years of formal preparation, but also through their local knowledge and "engine culture" spirit they succeeded where the government had repeatedly failed.

A second example is found in our long-standing efforts to immerse adults and children in the hard, but fruitful, work of inquiry and storytelling (Smith et al., 2000), where we have seen communities forge around journalism, creating stories of interest and concern for themselves and their Internet readers. This act of expression, facilitated by easy-to-use tools, led to an active debate over the content of their stories and, more importantly, the processes that they engage in as media producers. A collaborative editing process seems to help them develop a critical stance toward traditional media. As community participants challenge each other, they begin to understand the biases and critical thought processes that are the norm for professional journalists.

Rogoff (1995) describes the progression of community learning through the planes of apprenticeship, guided participation, and participatory appropriation. We propose to build technological affordances that serve in the role of emotional and inspirational mentor and that foster the creative and idiosyncratic connections to learning that help community members to progress through these planes.

CHALLENGE: Wear learning

"We need to brand math, and all learning, so that each morning, when youth stand in front of the mirror deciding who they will be that day, they always decide to wear learning." —Christine Ortiz, youth leader of Florida's "Truth" campaign

In 1998, a group of Florida teens was given authority and resources to launch a comprehensive campaign to change teen smoking behavior. The teens established a network of grassroots, youth-led community organizations, and organized a mass-media outreach which included television commercials (both professional and, more effective, 'unpolished' or homemade commercials), magazines advertisements, billboards, and flyers. After the first year of what was branded the "Truth" campaign, the program had a 92% recognition rate, equal to mega-brands like Nike. In the first three years of the program, smoking declined 47% among middle-school students, and 30% among high-school students. The proportion of students who had ever smoked also declined, from 44% to 32% in middle schools, and from 68% to 54% in high schools. A decline in smoking rates this large or even a decline at all, was unprecedented in Florida, and went against the upward trend in the rest of the nation.

Working with one of the youth leaders of the Truth campaign (now a student at MIT), we are exploring how the strategies underlying the successful campaign to influence youth attitudes about smoking could be adapted for a youth-led campaign to influence youth attitudes about learning. Studies have shown that youth attitudes about education (and about math in particular), and their perception of their own learning abilities (and, again, in math in particular), decline as students progress through K–12 schooling (see Figure 4, from US Department of Education, 2000). Could a youth-led campaign help transform these attitudes and perceptions?

The inherent differences between learning and smoking suggest that very different strategies might be needed. It is likely that long-lasting change in youth attitudes about learning will come only as a result of positive learning experiences, not simply positive messages about learning. But we believe that youth, particularly through the use of new media technologies, can play an important role in the effort to transform youth attitudes about learning. The Media Lab has already established a Young Activists Network (with chapters in a dozen communities) to support youth who are working to bring about change in their local communities. We plan to build on these experiences to engage youth in the effort to transform attitudes about learning.

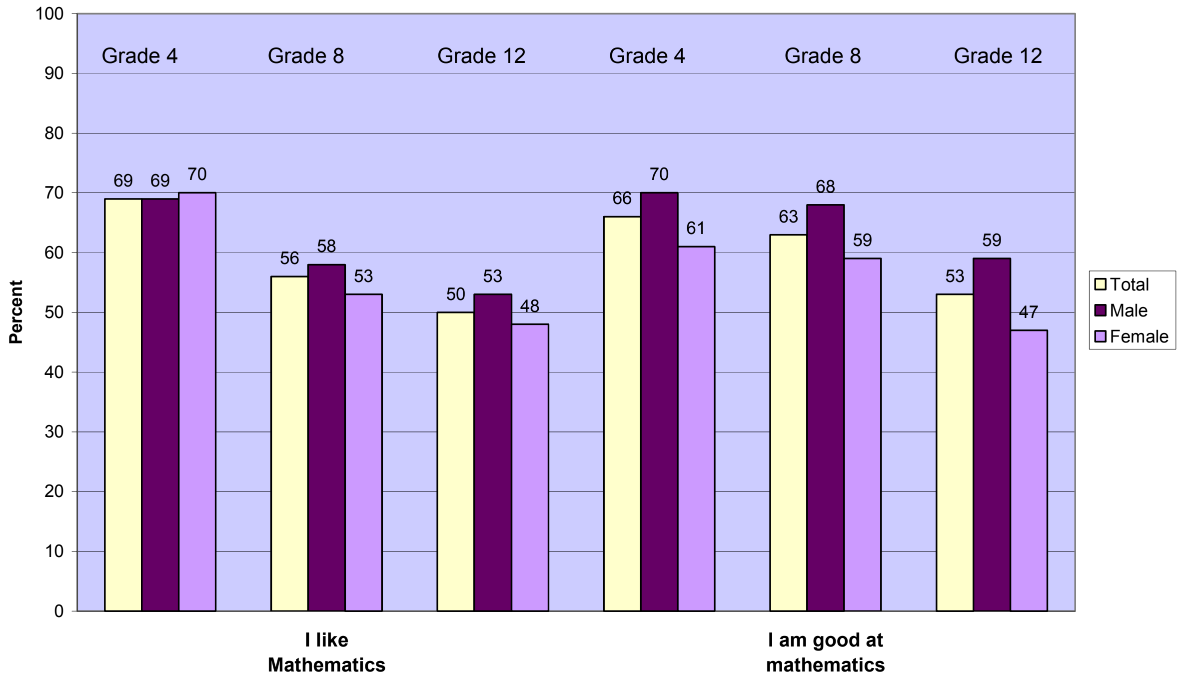

Figure 4: Student attitudes about mathematics decline from Grade 8 to Grade 12 (1996). Graph is from US Department of Education (2000).

OUTCOMES

Year 1: Design of prototype technologies and activities for all projects listed above, with participatory design input from appropriate user communities. For example, we will develop three new digital manipulatives for three different user communities; new sensor technologies, control electronics, and software for RoBallet activities; new designs for a next-generation community technology center; and prototypes of Music Games (each with explicit activity to achieve increased attention, memory, interior imagination, or collaboration). We will also organize a group of teens to work on the Wear Learning project.

Years 2–4: We will test all of the above prototypes in real-world testbeds, studying if and how prototypes support deep engagement and new conceptual understandings, and continually iterating our technologies and activities based on feedback from user communities. We will design and implement the first Garden for Creativity and Collaboration in the newly-constructed Media Lab building. Within Motivating Music project, we will develop a Haptic Playback system to provide guided, pedagogical experience of performing subtle gestural tasks. We will provide support and guidance for a selected group of youth in their efforts to transform youth attitudes about learning, which may include additional fund-raising from other sources. Based on all the above projects, we will continue our research on developing new theoretical frameworks for understanding deep engagement.

Years 5: We will work with Media Lab partners to broadly disseminate technologies and activities developed in Years 2–4 in some cases through commercialization in collaboration with corporate partners, in other cases through free distribution on the Internet. We will focus our research efforts on publication and dissemination of new frameworks for understanding deep engagement.

Partnerships, Outreach, and Dissemination

The Media Laboratory has an exceptional track record for getting its learning-related ideas and technologies out to the world in large numbers. The LEGO Mindstorms robotic construction kit, based on Media Lab research, has been used by millions of youth around the world. Software developed at the Media Lab (including StarLogo modeling software, Design-by-Numbers interactive-graphics software, and Pluto/HDL publishing software) is available for free online, and currently used in thousands of schools, community centers, and even senior centers. Fisher Price has announced plans to introduce, in the next year, a set of music-learning toys inspired by the Media Lab's Toy Symphony project, and the LEGO Company has recently announced a new product line, "Record & Play," which draws on its collaboration with the Media Lab's Tangible Media group.

The Media Lab has put its learning theories into practice by creating new educational contexts and settings—with particular focus on reaching youth from underserved populations. For example, the Media Lab co-founded the Computer Clubhouse project, a network of after-school learning centers for youth from low-income communities, with more than 80 sites in 15 countries, reaching more than 20,000 youth. With support from NSF, the Media Lab co-founded the Playful Invention and Exploration (PIE) Network, a network of museums that uses Media Lab technologies to create new types of public programs for children and families, integrating art, science, and engineering. Through its Digital Nations initiative, the Media Lab collaborates in a wide variety of international learning/education projects, with special focus on the developing world.

The Media Lab has also been very successful in sharing its ideas with the academic community through collaborations with other universities, publications in leading journals, and presentations at major conferences around the world. The senior personnel on this proposal have active collaborations with dozens of external institutions.

Recently, MIT announced its OpenCourseWare initiative, http://ocw.mit.edu, making MIT course materials freely available to the world. Although the full catalog of courses isn't expected to be available until 2007, the impact has already been significant both internationally and domestically. Following in this spirit, we propose not to limit the partnerships of the Center for Affective Learning to just a few selected institutions, but rather to open up what we do to the world, through a variety of new initiatives, detailed below. We aim to construct both online and face-to-face mechanisms ensuring significant exchange of people, ideas, and data, so that the Center functions as an international "hub" for research related to affect in learning.

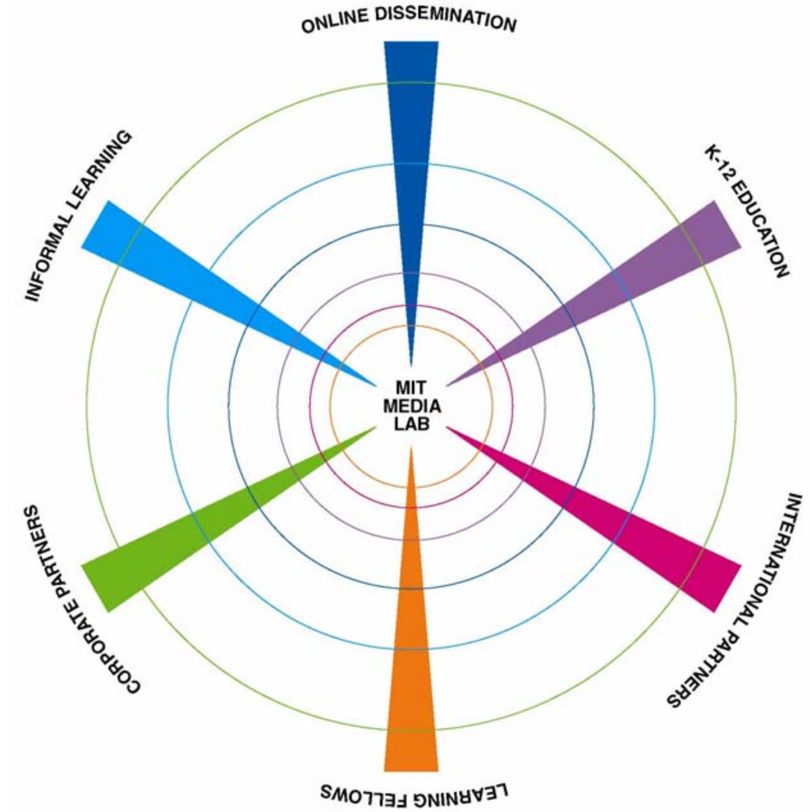

Figure 5: The Center for Affective Learning extends far beyond MIT.

- **Online dissemination**. The Center will distribute a variety of materials and resources free of charge via the Internet. The Center's colloquia and symposia will be Webcast and archived online, with indexing to make them more searchable. The Center will produce a special series of online discussions/debates among the world's leading learning and education researchers, together with experts from other fields who bring knowledge about affect (e.g., neuroscience researchers and others who study and build models of these phenomena). Efforts will be made to enable prototypes of hardware and software tools, such as new affect-sensing technologies, to be made available to researchers elsewhere. The Center will also make all of its courses available online through OpenCourseWare.

- **Fellowships**. One form of "reaching out" is "inviting in." The Center for Affective Learning will create a "Learning Fellows" program, analogous to Nieman Journalism Fellowship at Harvard, but designed for leading researchers and practitioners from the fields of learning and education. This program serves two purposes: (1) bringing new perspectives and competencies to the Center each year, to enrich and diversify the Center's ideas about learning and education; and (2) providing an organic way for ideas from the Center to get out into the world, since Fellows could naturally integrate ideas from the Center into their own local projects after returning home from the fellowship year. The Center will select four Learning Fellows each year, with priority given to candidates from under-represented groups. One of the priorities will be to use this as a vehicle for bringing in faculty experts on motivation, interest, and other affective aspects of cognition and learning. In addition to the Learning Fellows, the Center will also host visiting graduate students for periods of up to a year.

- **Informal learning**. The Center will work closely with the Media Lab's existing networks of Computer Clubhouses, community centers, museums, senior centers, and other informal-learning organizations. These collaborations provide not only a valuable public service, but also an authentic context for testing and evaluating new learning theories, strategies, and technologies developed at the Center. For projects aiming to engage the interests and passions of learners, informal-learning settings (where learners come by their own choice) offer a particularly important testbed.

- **K–12 education**. The Center will work closely with MIT's Teacher Education Program (TEP), which offers courses and student-teaching internships to MIT undergraduates who want to become K–12 teachers (particularly in math and science). Through this collaboration, the Center will create teacher-education materials based on ideas and technologies developed at the Center, test new ideas and technologies at Boston-area schools (leveraging TEP's strong relationships with local schools), and run summer workshops where K–12 educators can immerse themselves in the ideas and technologies developed at the Center.

- **Corporate partners**. The Media Lab has collaborative programs with many of the world's leading technology companies, such as Intel, Motorola, IBM, Microsoft, BT, Nokia, Sony, and Samsung, and it has established a strong infrastructure for technology transfer. Many Media Lab projects have been commercialized by sponsoring companies or by startup companies created by Media Lab alumni. We expect a similar pattern of technology transfer for projects developed at the Center.

- **International partners**. The Center will collaborate with international partners to explore learning issues in other cultural contexts. These partnerships will enable us to study which ideas transcend cultural specifics (and which do not). The Center will leverage the strong connections established by the Media Lab's Digital Nations consortium, which has collaborators in Mexico, Costa Rica, Panama, Colombia, and Brazil, including the Bradesco foundation, the Inttelmex foundation, INCAE, and SENACYT, as well as the international members of the Computer Clubhouse and PIE networks. The Media Lab additionally has a close partnership with MediaLabEurope, which brings expertise in the design of affective games, physical interfaces, analogical reasoning, and learning environments.

Evaluation and Assessment

Evaluation and assessment will be an ongoing process. We will continually adapt and revise our research plans and projects based on critiques and feedback from both internal and external sources. We will foster an environment of informal open critique, to ensure continual feedback from peers and colleagues. The routine operation of the Center will include weekly working lunches of the Center community, along with a weekly research seminar series to encourage and highlight progress and interdisciplinary exchange, as well as a variety of weekly project meetings. Additionally, we will subject our work to academic peer review.

In order to evaluate progress toward scientific outcomes, educational, outreach, and management goals, there will be regular meetings of the leadership team (PIs, Co-PIs, Deputy Director, and various advisors and project leaders) where goals are monitored and discussed. Additionally, there will be semi-annual program planning meetings with both internal and external partner organizations, where topic selection, project termination, and other decisions regarding resource allocation will be made. There will also be an annual research review by the Advisory Board coinciding with a briefing and written annual report to the NSF, providing oversight, guidance, and help with prioritization.

Advisory board. The following individuals have agreed to be a part of the Advisory Board for the Center for Affective Learning: **Michael Arbib**, Fletcher Jones Professor and Chairman of Computer Science, and also Professor of Neuroscience, Biomedical Engineering, Electrical Engineering, and Psychology at the University of Southern California USC; **John Seely Brown**, Director Emeritus of Xerox PARC, and Visiting Scholar at the Annenberg Center at University of Southern California; **Red Burns**, Founder and Chair of the Interactive Telecommunications Program at New York University; **Howard Gardner**, John H. and Elisabeth A. Hobbs Professor in Cognition and Education at the Harvard Graduate School of Education and Adjunct Professor of Psychology at Harvard University, as well as Adjunct Professor of Neurology at the Boston University School of Medicine; **Alan Kay**, Senior Fellow at Hewlett Packard Labs, and President of Viewpoints Research Institute; **Walter Massey**, President of Morehouse College and former Director of the National Science Foundation, **Bill Wulf**, President of the National Academy of Engineering.

Management Plan

The Center for Affective Learning will be directed by Professor Picard, who will be responsible for the overall supervision and coordination of activities at MIT. Picard will also represent the Center as a member of the SLC National Network. Both MIT and the Media Lab will provide a supporting institutional context for the Center, providing services such as financial oversight and controls, facilities, network services, management of intellectual property, and publications. The co-PI's and other senior personnel will have responsibility for supervising MIT and visiting graduate- and undergraduate-student researchers in all aspects of the research, from software development to data collection, analysis, and documentation.

- **Coordination of data collection and analyses.** The co-PIs, together with graduate student researchers, will be responsible for designing and implementing research activities. Each co-PI will develop appropriate research procedures and will obtain prior approval from the MIT Committee on the Use of Humans as Experimental Subjects for research involving human subjects. Walter Bender will oversee this process. Henry Holtzman will oversee the creation of a central repository for indexed field notes, video logs, log files and other artifacts so that the proposed analyses of data can be coordinated and shared beyond MIT (as appropriate).
- **Reporting mechanisms.** Picard will be responsible for filing annual progress reports to the NSF. She will be assisted by the co-PIs in collecting information about ongoing project activities. Research meetings with co-PIs will be scheduled on a quarterly basis with conference calls and face-to-face meetings occurring often more frequently as collaborations dictate. To facilitate sharing of information among project members on an ongoing basis, we will make use of various online collaboration technologies (including mailing lists and "wikis").
- **Education.** The Media Lab has been a leading innovator in curriculum development and in advancing an "atelier" model for learning, emphasizing studio-style workshops in which students work on a series of design projects with rapid design-critique-redesign cycles. In our experiences at the Media Lab over the past 18 years, we have found that atelier-style courses are a particularly effective way for undergraduate and graduate students to explore new ideas and technologies, integrate skills from different technical disciplines, and develop as creative designers and inventive thinkers. The Media Lab has also been a leader in introducing this style of learning and education to broader audiences and contexts, from kindergarten students to senior citizens, from school classrooms to community centers, from inner-city US neighborhoods to rural communities in the developing world. The Center for Affective Learning will continue and extend this tradition, integrating its research findings into new learning/education initiatives inside MIT and beyond.

- **Technology transfer.** Industrial technology transfer will be based around the various Media Lab sponsor programs, which engage approximately 100 companies. These companies, which range from a variety of industry sectors (from electronics to telecommunication to manufacturing to service industries), participate through regular visits and semi-annual review meetings. They provide complementary funding, a vehicle for potential commercialization of results, and also insight into the practical development of fundamental research. In addition, the Center will engage with a number of educational institutions and foundations worldwide through its collaboration with the Media Lab's Digital Nations program.
- **Management of hardware/software development**. The co-PIs and senior personnel will oversee and coordinate hardware and software development. A version of "Source Forge," which has been modified to meet the internal requirements of the Media Lab, will be used to coordinate research activities. It utilizes a shared-file server at MIT to publish and archive system releases during testing phases of a project, and as a bug-reporting system to keep track of problems and feature requests. The Media Lab has an excellent information technology infrastructure, including state-of-the-art network security and a nightly server backup system.
- **Intellectual property.** The Media Lab has a liberal intellectual property policy that is unique for university research laboratories: In order to foster open discourse among researchers and sponsors from industry and governments, sponsors have the right to royalty-free license to *any* technologies developed at the Lab. In addition, Media Lab faculty, students, and alumni have been pioneers in the Open Source movement, and the vast majority of the intellectual property created at the Lab is put directly into the public domain.
- **Center administration.** The MIT Grants & Contracts Office has procedures in place to administer subcontracts and handle the invoicing procedures. The Media Lab's chief finance officer will provide direct fiduciary oversight. The Media Lab director's office will provide oversight of the contract and any intellectual property issues associated with the Center. These efforts will be coordinated with MIT's Office of Sponsored Research and Technology Licensing Office. Lab space, office space, and furniture will be provided by university administration and be managed by the Media Lab's facilities manager. The Media Lab's Network and Computing Systems will provide support for the Center. A program manager will be in charge of all administrative issues and logistics for project-wide meetings and events at MIT. The administrator for the MIT Program in Media Arts and Science will provide support for the proposed Fellows program.
- **Diversity.** MIT is committed to the principle of equal opportunity in education and employment. The Institute does not discriminate against individuals on the basis of race, color, sex, sexual orientation, religion, disability, age, veteran status, ancestry, or national or ethnic origin in the administration of its educational policies, admissions policies, employment policies, scholarship and loan programs, and other Institute-administered programs and activities. Additionally, many of our outreach initiatives (especially the Computer Clubhouses and the Teacher Education Program) target under-represented minorities.
- **Sustainability.** The plan for long-term sustainability of the Center is based upon the Media Lab's nearly 20 years of success in combining industry, government, and foundation funding. We anticipate that over the lifetime of the Center, we will have "hardened" the faculty positions with chairs, migrated a majority of student fellowships to industry support, and obtained out-reach operating costs from international, national, and local foundations.

Prior NSF Support received by PIs and Co-PIs

Rosalind W. Picard
NSF ROLE award number REC-0087768
Amount & Period of Support: $947,008 for 36 months, 11/15/2000 through 10/31/2003
PI: Rosalind W. Picard, co-PIs: Justine Cassell, Rob Reilly, Barry Kort.
Title: The Role of Emotion in Propelling the Science, Math, Engineering, and Technology Process

Summary of Results: This research has advanced a new theory and developed new algorithms and tools for non-invasively sensing, recognizing, and responding to a child's emotional state in a computer learning situation. Results include: (1) developing a system to gather and synchronize data from multiple sensors, including two custom sensors (BlueEyes Camera and TekScan Chair pressure sensor); (2) crafting a computer learning interaction that reliably elicited both boredom and different levels of interest, and collecting five channels of synchronized data from twenty children engaging in this interaction; (3) developing, through iterative work with teachers, a set of emotion labels that could be reliably and meaningfully attached to this data, and having the data coded with these labels: "high interest," "medium interest," "low interest," "taking a break," and "bored;" and (4) developing and testing automated pattern recognition/machine learning algorithms for enabling the computer to infer these labels, and achieving highly significant recognition rates on upper facial expressions and on shifts in posture related to affective state; thus, showing that machines can recognize, with significantly higher than chance probabilities, indications of the learner's affective state related to interest and boredom.

In parallel with this effort, we developed tools for building agents that would respond to the learner's affective state. We defined, designed, and tested new "relational agents," agents capable of building long-term social-emotional relationships with people. We conducted a 99-person one-month test of the first relational agent, where subjects were split into three groups, all doing the same task (1/3 with no agent, 1/3 with the relational agent, 1/3 with the same agent but without its relational skills) and found the task outcome improved in all cases, while a "bond" rating toward the agent was significantly higher in the relational case. Thus, this work showed that certain relational skills could be automated, resulting in people reporting that the agent cared more about them, was more likeable, showed more respect, and earned more of their trust than the non-relational agent. People interacting with the relational agent were also significantly more likely to want to continue interacting with that agent. The significant difference in people's reports held at both times of evaluation: day 7 and day 27, showing that the improvement was sustained.

Publications: Aist et al., 2002a,b; Bickmore, 2001; Bickmore, 2002; Bickmore and Cassell, 2002; Bickmore and Picard, 2002; Bickmore, 2003; Bickmore and Cassell, 2003; Bickmore and Picard 2003; Cassell, et al., 2001; Kapoor and Picard, 2001; Kapoor et al., 2001; Kapoor 2002; Kapoor and Picard, 2002; Kapoor et al., 2003; Kort et al., 2001a,b; Kort and Reilly, 2001; Kort and Reilly, 2002a–e; Kort et al., 2002; Kort and Reilly, 2003a-c; Kort and Reilly, 2004a,b.; Liu and Picard, 2003; Mota, 2002; Mota and Picard, 2003; Picard, 2003.

Available databases and other products: The sensors and algorithms developed in this research are available, and have already been adopted by University of Memphis, NASA, and BBN for use in intelligent tutoring systems projects.

Mitchel Resnick
NSF Informal Science Education award number ESI-0087813
Amount & Period of Support: $3,134,757 (including subcontracts) March 2001–February 2004
PI: Mitchel Resnick, co-PIs: Bakhtiar Mikhak, Mike Petrich, Natalie Rusk, Karen Wilkinson
Title: The PIE Network: Promoting Science Inquiry and Engineering through Playful Invention and Exploration with New Digital Technologies

Summary of Results: The Playful Invention and Exploration (PIE) Network, a collaboration of the MIT Media Laboratory and six museums, has developed a new generation of events and public programs integrating art, science, and technology, linking the use of physical materials with digital technology in creative inquiry and inventive exploration.

PIE museums have integrated new MIT educational research and technologies (such as "programmable bricks") into their ongoing public programs. Museums have organized MindFest events, at which young people, educators, artists, engineers, hobbyists, and researchers come together to collaborate on invention projects and share ideas with one another. Museums have also organized a series of Make-Your-Own Workshops for families (with a focus on those with children ages 8–14). Workshop themes include: Make-Your-Own Kinetic Sculptures, Make-Your-Own Mood Meter, and Make-Your-Own Programmable Jewelry.

The PIE Network also contributed to the development of the Lemelson Center's *Invention at Play* exhibition, which opened in 2002 at the National Museum of American History and is now traveling to nine science centers and museums nationally, expecting to reach one million visitors. The exhibition includes a 10-minute video highlighting Media Lab learning research.

Publications include: Borovoy et al., 2001; Resnick, 2001; Resnick, 2002; Resnick, 2003; Zuckerman and Resnick, 2003a; Zuckerman and Resnick, 2003b.