Affective Medicine: Technology with Emotional Intelligence

Rosalind W. Picard

MIT Media Laboratory; 20 Ames Street; Cambridge, MA 02139, USA; http://www.media.mit.edu/affect

Abstract. For a long time people have kept emotions out of the deliberate tools of medicine and science; scientists, physicians, and patients have often felt and sometimes expressed emotion, but no tools could sense, measure, and respond to their affective information. A series of recent studies indicates that emotions, particularly stress, anger, and depression, are important factors with serious and significant implications for health. This paper highlights research at the MIT Media Lab aimed at giving computers the ability to comfortably sense, recognize, and respond to certain aspects of human emotion, especially affective states such as frustration, confusion, interest, stress, anger, and joy. Examples of recently developed systems are shown, including computer systems that are wearable and computers that respond to people with a kind of active listening, empathy, and sympathy. Results are reported for computer recognition of emotion, for teaching affective skills to autistics, and for having computers help users manage emotions such as frustration.

1 Introduction: Frustration, Irritation, Stress and Health

Perhaps the most common emotions people feel in interacting with today's technology are frustration, irritation and other feelings related to stress. We've been doing experiments in our lab where we bring in people, give them a task on the computer that mildly or strongly frustrates them, and measure how they behave. Our aim is to try to teach the computer how to recognize when the user is frustrated, irritated, annoyed, stressed or otherwise in some significant emotional state, and then to equip the computer so that it can do a better job of serving people, ideally not causing them so much stress. It's recently become a joke in the lab when some piece of equipment fails or causes aggravation – "is this one of Picard's affective computing experiments designed to irritate me?"

What could technology do if it could sense that the user is frustrated or otherwise in some unusual emotional state? Could the system change or be changed, so as to reduce frustration in the future, or could it help the user then and there to feel less stress? Our research has focused on both approaches: identifying components of computer interfaces that could be improved by designers, as well as having the computer help users manage strong negative emotions better. This is all part of our effort in "affective computing," computing that relates to, arises from, or deliberately influences emotion [18]. This article will focus the topic further -- on some recent findings about emotion and medicine -- together with examples of new affective technology we are developing that has potentially interesting and important implications for health.

Stress is increasingly recognized as a medical problem. A recent Blue Cross Blue Shield survey in New England cited the number one health concern of members in this part of the United States to be stress – rated above cancer, AIDS, high blood pressure, and other medical conditions. Dan Goleman, in Chapter 11 of his book Emotional Intelligence [7], cites a number of studies pointing to important roles in health and medicine for emotions – particularly states of stress, anxiety, chronic anger, and depression. Following are just a few of the examples that he includes:

> *Stress/distress:* Studies of the physical manifestations of stress reveal many measurable changes caused by stress in the human body, influencing not only immune system functioning but also heart rate variability, blood pressure, and other important bodily functions. Several studies have been conducted examining the impact of stress on immunity. For example, Sheldon Cohen, a psychologist at Carnegie-Mellon University exposed people to a cold virus after assessing how much stress they were experiencing in their lives. Of course, a robust immune system usually resists a virus, so mere exposure doesn't mean you will get sick. Cohen found that 27% of the low-stress subjects came down with a cold while 47% of the high-stress people came down with the cold. In another study of married couples who kept daily logs of hassles and upsetting events, a strong pattern emerged: three or four days after an especially intense marital fight or other upset, they came down with a cold or upper-respiratory infection.
>
> *Depression:* In work cited by James Strain, where 100 patients received bone marrow transplants, a follow-up study was conducted of the 13 who had been depressed vs. the other 87. Of the 13 who had depression, only 1 was alive a year later; of the other 87, 34 were still alive two years later. Another study, by Howard Burton et al., associated early death of dialysis patients with depression: depression was found to be a stronger predictor of death than any medical sign. Similarly, heightened risk of death from heart disease has been correlated with an ongoing sense of despair and hopelessness.   The effect of depression on heart attack survivors is as great as that of major medical risks such as left ventricular dysfunction or a history of previous heart attacks.
>
> *Anger:* Dr. Redford Williams at Duke University found that physicians who scored highest on tests of hostility while they were in medical school were seven times as likely to have died by the age of fifty as those who scored low on the hostility tests – their tendency to get angry was found to be a stronger predictor of early death than were factors such as high blood pressure, high cholesterol, and smoking.   Findings by Dr. John Barefoot at the University of North Carolina show that scores on a test of hostility correlate with the extent and severity of coronary artery disease in heart patients undergoing angiography.

The studies above are but a few of the many that reveal emotion to be a measurably important health factor. Note that none of the studies show that emotions *cause* medical illness – rather they contribute to decreases in immune system functioning and to other physical factors that may significantly prolong or exacerbate an illness. It is foolish to think

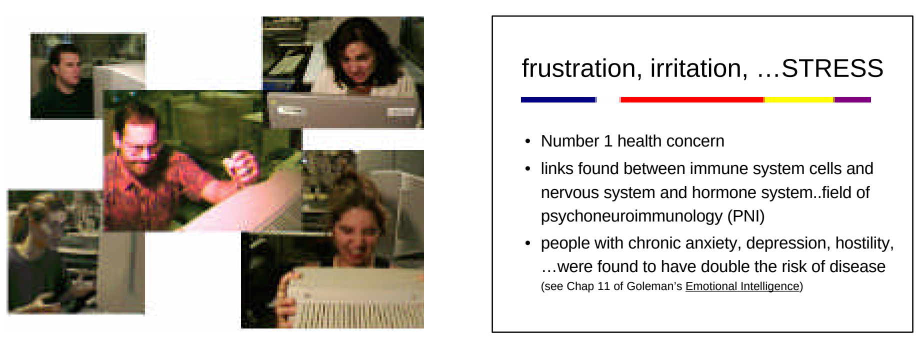

Figure 1. Stress is a significant factor in health; it arises in many forms when interacting with technology.

that "positive thinking" or "making yourself happy" can prevent all illness; it is also foolish to continue to think that emotions have no significant effect on health – the truth appears to lie somewhere in the middle, with emotions playing not the only role, but an important measurable role that has typically been ignored.

In the Affective Computing group at MIT, we are particularly interested in the intelligent handling of affective states commonly expressed around computer systems: frustration, confusion, disliking, liking, interest, boredom, fear, distress, and joy. Computers and other forms of technology are interacting with people in more ways than ever before – beyond desktop, laptop, and palmtop, technology is now embedded into appliances, clothing, jewelry, implants, and even pills we can swallow. With all these new forms, technology has the opportunity to detect physical and physiological expressions of many human emotional states. With additional sensing and processing, the expressions of emotional state can be associated with other events – such as what the person is doing when they get angry or stressed, what else is happening in their body concurrent with episodes of depression, (perhaps related to their heart functioning or their physical activity level) or what the interface (if the person is interacting with one) may have just done [19].

The rest of this paper is divided into four areas: (1) Sensors that enable the user to communicate information related to emotion in a way that is physically and psychologically comfortable; (2) Progress in computer recognition of emotion; (3) Tools for helping people learn affective skills, including a system for autistic kids; (4) Respectfully handling emotions, such as reducing user frustration.

2 Comfortable sensing of signals

Emotions often involve both thinking and feeling – both cognitively experienced events and physical changes in the body. Although there is no technology that can truly read your thoughts, there are a growing number of sensors that can capture various physical manifestations of emotion -- video recordings of facial expressions and posture or gesture changes, microphone recordings of vocal inflection changes, skin-surface sensing of muscle tension, heart-rate variability, skin conductivity, blood-glucose levels, and other bodily changes, and (if invasiveness is allowed), swallow-able or implant-able sensors or means of capturing bodily fluids for analysis. These are just a few of a growing number of possibilities.

Our research efforts include building tools to facilitate multiple forms of emotion sensing, not to force this on anyone, but to allow for a larger space of possibilities for those who want to communicate and better understand affective information. The tools include new hardware and software that we have developed to enable certain machines not only to

Emotions give rise to changes that can be sensed

	Face
Distance	Voice
Sensing:	Posture
	Gestures, movement, behavior
	Skin conductivity
	Pupillary dilation
Up-close	Respiration, heart rate, pulse
Sensing:	Temperature
	Blood pressure

Figure 2. Emotion influences many changes in the body. These patterns of change can be sensed by various new wear-able or swallow-able technologies.

receive emotional expression, but also to recognize meaningful patterns of emotional expression. In particular, we have integrated several physiological sensors into clothing and jewelry – a blood volume pressure sensor in an earring, skin conductivity in a shoe and a glove, respiration in a sports bra, and more [19] (See some examples in Figure 2). These sensors communicate with new wearable computers that can control peripherals such as a wearable music player/DJ [8] or a wearable camera [10]. For example, the wearable camera system we built saves video based on your arousal response tagging the data not just with the usual time stamp, but also with information about whether or not it was exciting to you, as indicated by patterns it detects in your skin conductivity [10]. The same system could potentially be modified to detect and communicate health-related variables to you and your physician, perhaps for monitoring and analyzing patterns of stress, anger, or depression.

With any wearable system there are design issues regarding not only what is to be sensed, but also how the sensing system can be made comfortable and robust to noise that arises from activity unrelated to the signal being measured. The key source of noise when measuring emotion from ambulatory patients is artifacts that arise from physical activity. Heart rate, for example, can increase significantly with physical exertion or with sneezing, as well as with anger and other affective states. Inferring the source of a change is easier if you can independently detect the change – such as via context sensors that indicate the person's movement or activity.

One of the physiological sensing systems we built that is robust to motion artifacts is the "Conductor's Jacket," one version of which is shown in Figure 3. This highly expressive wearable system, created by Teresa Marrin, associates patterns of muscle tension and breathing with expressive gestures that the conductor uses to shape the music. Seven electromyogram (EMG) and one respiration sensor are included in the version shown here. The EMG's are attached with custom-fit elastics sewn into the shirt, so that they remain snug without strong adhesives, and yet do not move as the arms are moved. This wearable system was designed first to measure how professional and student conductors naturally communicate expressive information to an orchestra. After analyzing real conducting data from six subjects, Marrin found around thirty significant expressive features (largely related to muscle tension changes that signaled interesting musical events) [16]. She has subsequently developed a version of the jacket that transforms natural expressive gestures of the wearer into real-time expressive shaping of MIDI music [17]. A professional conductor, Marrin is currently using the jacket both for live performance and for helping educate student conductors, providing precise feedback on timing, tension, and other important aspects of expressive technique.

In addition to the goal of making wearable sensing devices robust and physically comfortable, we have been concerned about the psychological factor – how do these sensors feel from a personal comfort standpoint and within a social setting? We do not believe there

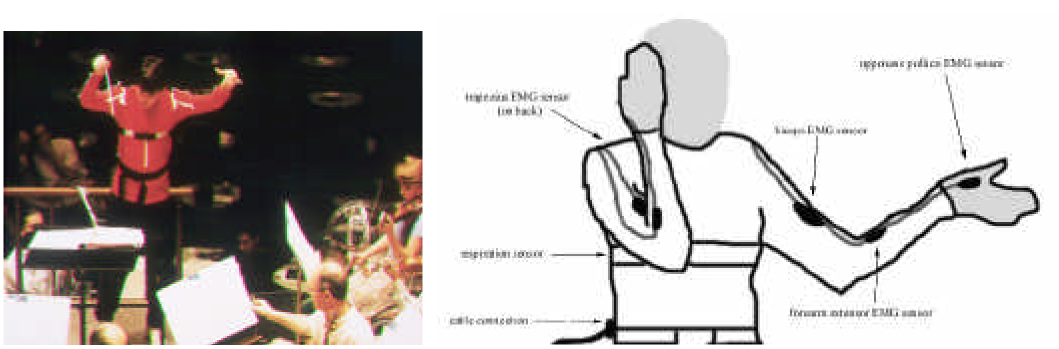

Figure 3. The "Conductor's Jacket" used both for gathering physiological changes related to expression and for controlling MIDI output.

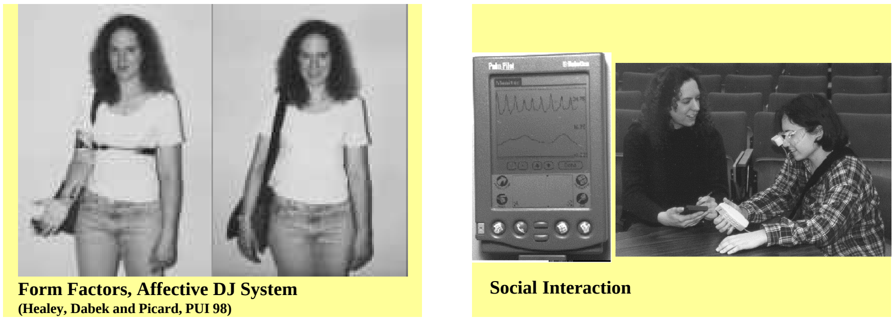

Figure 4. Wearable systems can be increasingly hidden in ordinary clothing; their form factor can be modified to not impair social interaction.

is a one-size-fits-all answer, but rather we find that people like to exercise choices; in many cases, these choices include hiding the sensors, so that they are not visible (although a notable exception was one of the professional conductors we worked with – we took him a sensor jacket with almost everything hidden under the tuxedo, and he asked us to redo it "in red, with the wires in silver on the outside for all to see.") We found that when we integrated sensors with the "lizzy" wearable (Figure 4, at right) that the private eye output display was an impediment to social interaction, and the input device (chording keyboard) took too long for subjects to learn. Consequently we modified a Palm Pilot to serve as an input/output device to the wearable, allowing the Palm to display physiological signals and receive annotations for them. Although a palm device is not hands-free like a heads-up display, it is more comfortable for many people in social settings.

Emotions modulate not only the modes shown in Figure 2, but also many others, including hormone and neurotransmitter levels. The latter are currently not easy to sense without drawing saliva or blood or using other invasive procedures. Presently, none of these procedures provides instantaneous wireless access to the changing levels. However, new implant-able and swallow-able sensors are being developed by many researchers, exploiting strides in nanoscale technology, and giving access to internal bodily signals previously unavailable in real time (e.g., Figure 5, courtesy of Prof. Scott Manalis at MIT).

Sometimes, social-psychological concerns such as privacy make one form of sensing preferable to another. For example, although one person might be comfortable communicating facial expressions to a computer using a video camera, another might be concerned about the identifying information that the camera would see. In distance learning and in intelligent tutoring systems, there is an opportunity for the student to transmit signals

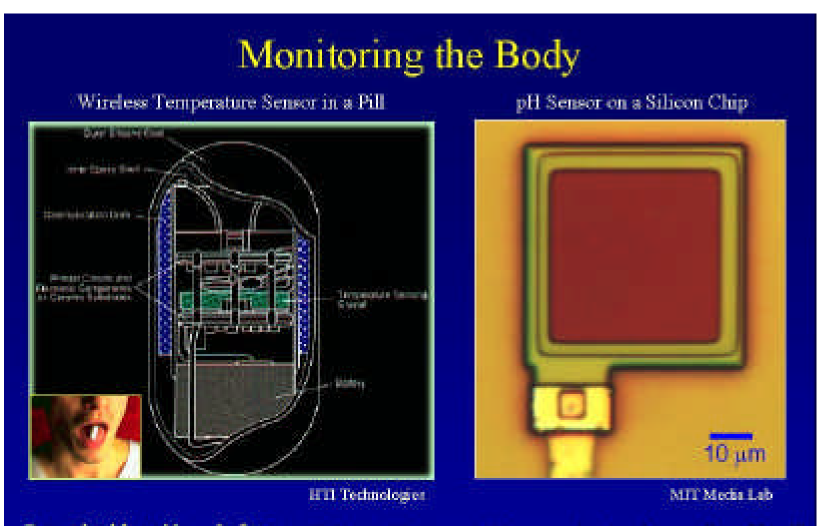

Figure 5. New forms of sensing enable real-time wireless readout of internal bodily signals such as temperature or pH.

such as confusion or interest in real-time, without having to stop and click on anything that interrupts their attention [22], and without giving away their identity. One means of accomplishing this is via a wearable sensing system designed and built by research assistant Jocelyn Scheirer, the “expression glasses” described in the next section.

3 Computer recognition of emotional expression

One of the wearable-computing platforms we built includes a small A-to-D with eight channels for physiological sensing. We have developed algorithms that run on the wearable system, extract features from the physiological signals, and relate these to a deliberately expressed emotion. Short segments of four physiological signals for two emotions are shown in Figure 6. Although the segments here look different for each emotion, this was not always the case; in general, the variations within the same emotion from day to day exceeded the variations in different emotions on the same day. Using a variety of methods of pattern recognition and baselining, we have obtained recent results of 81% recognition accuracy in selecting which of eight emotions was expressed by an actress, given 30 days of data, eight emotions per day, and features of the four signals: respiration, blood volume pressure, skin conductivity, and muscle tension. (See Healey and Picard (1998) and Vyzas and Picard (1999) for details of the data collection and the recognition algorithms [9], [25].) The eight emotions investigated were: neutral, hatred, anger, romantic love, platonic love, joy, and reverence. These are the best known results to date for emotion recognition from physiology, and they lie between machine recognition results of affect from speech and of affect from facial expressions.

It should be noted that these results are for a single user, and they are obtained by a forced selection of one of the eight categories; hence, these results are comparable to recognition results in the early days of speech recognition, when the system was re-trained for each speaker, and it knew that the person was speaking one of eight words, although there could be variation in how the person spoke the words from day to day. Much more work remains to be done to understand individual differences as well as differences that depend on context – whether developmental, social, or cultural. I expect that, like research in speech recognition, this work will gradually expand to be able to handle speakers from different cultures, of different ages, speaking (or expressing) continuously, in a variety of environments.

Figure 6 (right) shows a computer task and data-gathering system we designed that was intended to induce negative stress and collect data synchronized with the stress-eliciting events [21]. We gave the user a goal with incentive: race through the task as quickly as possible, obtain the best score (a mix of accuracy and efficiency) and win a $100 prize. Along

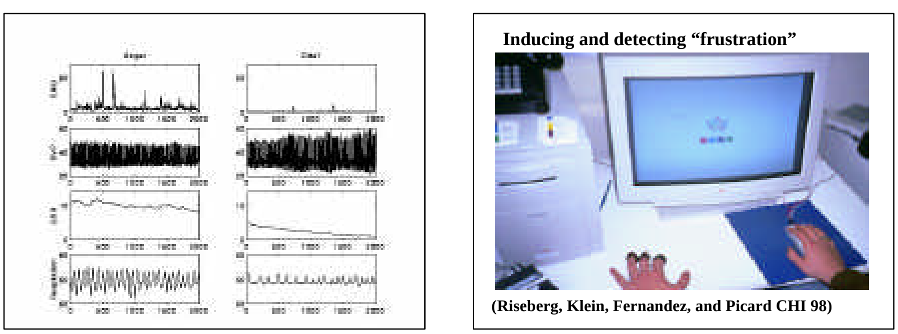

Figure 6. Examples of four physiological signals sensed during different emotional states.

the way, we had the system freeze up as if the mouse was not working, delaying their progress. We continuously measured two physiological signals—skin conductivity and blood volume pressure—then compared patterns in these signals when all was going smoothly vs. during the episodes of unexpected delays. Although we cannot determine whether these episodes corresponded to true feelings of frustration or non-frustration, we did find that in 21 out of 24 subjects, the patterns detected by our Hidden-Markov-Model based approach were able to significantly discriminate these two kinds of episodes [5]. However, the recognition results were still far from perfect, indicating that although this information is helpful, it must be combined with other signals for a more confident decision.

Stress is sometimes a by-product of feelings such as confusion, which a person may choose to communicate by furrowing his or her brow. The furrowing of the muscle can be detected by a camera if lighting and head position is carefully restricted (otherwise current computer vision techniques are inadequate) but these restrictions, coupled with the recording of identity, can make some subjects uncomfortable. An alternative sensor to a camera is a pair of wearable "expression glasses" (Figure 7) that senses changes in facial muscles [22]. These glasses have a small point of contact with the brow, but otherwise are considered by some users to be less obtrusive than a camera in that the glasses offer privacy, robustness to lighting changes, and the ability to move around freely without having to stay in a fixed position relative to a camera. The expression glasses can be used *while concentrating* on a task – the wearer does not have to stop and think about how to communicate a facial expression. The glasses can be activated either unconsciously or consciously. People are free to make false expressions, or to have a "poker face" to mask true confusion if they do not want to communicate their true feelings, but if they want to communicate them, the glasses offer a virtually effortless way to do so.

Why wear expression glasses, instead of raising your hand or pushing a button to say you're interested or confused, as was implemented decades earlier [23] by Sheridan and his colleagues? The answer is not that there should be one *or* the other; both kinds of feedback offer advantages. Sometimes people miss the subtlety of this point: affect is continuously communicated while you are doing just about anything. When you pick up a pen, you tend to do so very differently when you are angry vs. when you are joyful. When you watch somebody, your eyes behave differently if you are interested than if you are bored. *As you listen* to a conversation or a lecture, your expression gives the speaker feedback, unless, of course, you put on a poker face. In contrast, if you have to think about pushing a button to communicate your feelings, or to raise your hand to say you're confused, then you have to interrupt your concentration to take such an action. Self-report is important, but it is no substitute for the natural channels of largely non-verbal communication that humans use *concurrently* while engaged in conversation, learning, and other activities.

One of the important domains for analyzing stress is with drivers, especially drivers

Figure 7. Expression glasses sense facial muscle changes and detect furrowing of the brow, a signal sometimes used to communicate confusion.

trying to do more than drive. We have conducted experiments measuring the impact of low and high cognitive load tasks on drivers talking over a telephone headset while driving in a simulator. We placed drivers under different load conditions (fast or moderate speed, and fast or slow questioning with simple arithmetic problems like "12 + 14") while otherwise keeping the driving task the same. The drivers were occasionally exposed to signs labeled "brake" or "continue" and were instructed to brake as soon as they saw the brake message. Most drivers braked within 0.7-1.4 seconds after the message; however, there were a number of incidents where braking took place 1.5-3.5 seconds after the brake message, or not at all. In almost all of the latter cases, the subject was talking on the phone. On average, the drivers talking on the phone had reaction times to brake messages that were 10% slower than when they were not on the phone; more importantly, the variance in their braking times was four times higher – suggesting that although delayed reactions were infrequent, when delays happened they could be very large and potentially dangerous. The fact that they didn't happen often could furthermore create a false sense of security. Although physiological data gathered in these experiments was limited, our analysis indicated a potential for recognizing patterns that might indicate whether or not a driver was likely to respond with a slowed reaction or not [24].

We are beginning to analyze affect in speech, an area in which humans perform at only about 60% accuracy (on roughly eight emotion categories, when the content of the speech is obscured). Our initial focus is on speech from drivers, taken from the experiment above, examining if the driver's vocal characteristic under different load conditions shows reliable indications of stress [6]. As manufacturers put more gadgets in cars, such as talking navigation systems and restaurant guides or grocery reminders triggered by GPS-sensed location, there is increased potential that the driver might be interrupted at a stressful time that could diminish safety. Another human passenger would be able to sense if the timing was good or not and make a safer decision about interrupting the driver; however, the systems being put into cars are currently oblivious to these factors. If the driver is conversing with one of the car systems, he or she may be distracted at a bad time, compromising safety. We are trying to give the system the ability to sense stress pattern changes in the driver's speech (as well as responsiveness patterns in the driver's behavior, per above) so that the car can be more sensitive to the safety factor.

Our research is developing means of recognition of physiological patterns related to stress in many different natural environments. We have recently moved outside the world of simulators and equipped a car to examine driver behavior features joint with physiological information. One such sensor set-up is shown in Figure 8 (right). In a recent set of experiments we induced stress in a dozen drivers by having them drive around Boston under four stress-eliciting conditions while we recorded electrocardiogram, skin conductivity,

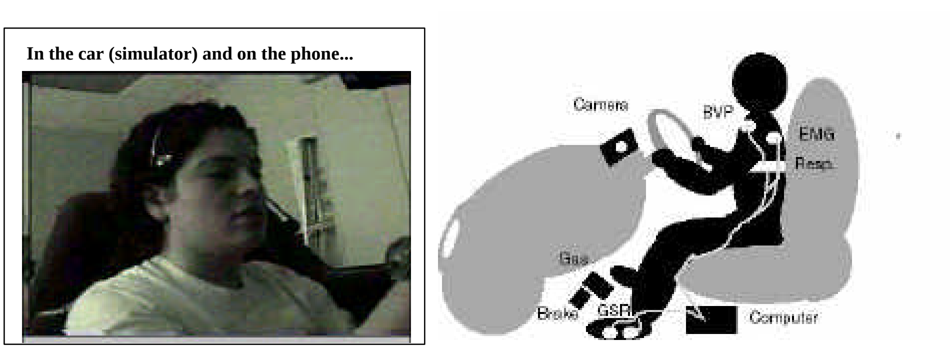

Figure 8. Stress was measured from vocal intonation and from driving behavior when drivers were given tasks over a phone headset (left). We designed and integrated a physiological sensing system into a Volvo for measuring bodily signals related to stress among Boston drivers out on the road (right).

respiration, blood-volume pressure, and electromyogram signals, together with video and other information about the driver's behavior and context. Knowing how stressed somebody truly became is hard to assess; therefore, we used three different methods of assessment: self-report of the driver, driving condition (rest, busy city, easy highway, tolls/turn-around), and third-party coding of complexity level based on number of events each minute during the driving situation. Our analysis of patterns of driver physiology, led to an average stress recognition rate of from 89-96% accuracy depending on which of three methods was used for labeling the "true" stress level of each subject [11].

An increasing amount of human-human communication takes place through machines. In many cases it would be helpful if the machine would simply facilitate the transmission of affective cues. An example of a system designed to expand human-human communication capabilities via computer is the TouchPhone, developed by Jocelyn Scheirer in our lab (Figure 5). The TouchPhone augments regular voice communication with pressure information indicating how tightly the speaker is holding the phone. For example, if you routinely talk to an elderly parent by phone, this would enable you to not only hear their voice, but also to see how they were holding the phone: Is it the same as most days? Or today does their grip seem weaker, tense, or more fidgety? The pressure is continuously mapped to a color seen by the person on the other side – calibrated to blue if light pressure is applied – and to red if strong pressure. The computer performs no interpretation of this signal; the color signal is simply transmitted to the conversational partner as an additional low bit-rate channel of information.

I met with four of my students for four hours of TouchPhone conversations and the results, while anecdotal, were interesting and were consistent with experiences we have had with other emotion-communication technologies. I found that each of the four students had a nearly unique color pressure pattern, which was distracting until I moved the pattern into my periphery where it became ambient, adding a flavor of background rhythm to the conversation. For one student, the pattern changed very slowly, becoming stable red when I started asking some research questions. I thought nothing of it, because he could have simply been squeezing the phone more tightly by shifting his position. However, even though he knew that I could not tell his feelings from the color, he expressed to me that he wasn't trying to squeeze it tighter at all and he thought it was red because he was stressed about a question I asked him. The student was a non-expressive male engineer who had never revealed such signs of stress to me in the years of conversations we had had prior to this TouchPhone conversation. The technology thus facilitated opening up a greater range of emotional communication – by his choice – it did not impose this, but simply made it easier for him. The color did not give away how any of the students was truly feeling. However, the system provided a new channel of non-verbal communication that, in turn, could and did sometimes open up a new line of verbal communication.

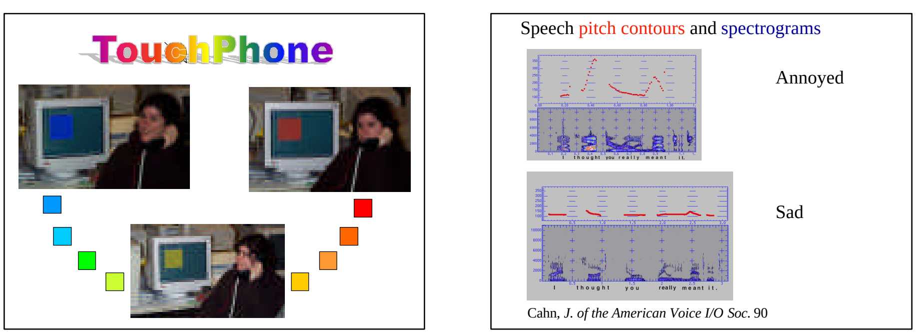

Figure 9. System that senses how phone is being held (left.) Examples of intonation changes in annoyed and sad speech (right).

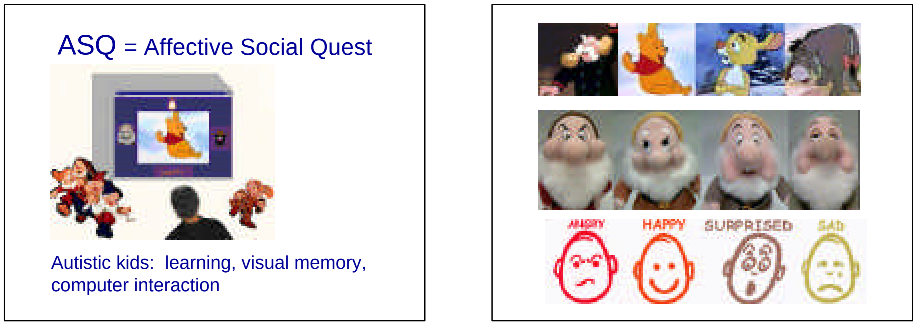

Figure 10. The “ASQ” Computer system shows video clips to autistic kids and prompts them to choose the stuffed dwarf that expresses the emotion appropriate to the video scene. The system senses the child’s response and rewards accordingly.

4 Helping build human emotional skills

Computers just “don’t get it” when it comes to practicing many of the social-emotion skills that most of us take for granted. Although autism is a complex disorder, and some of the comments here will not apply to all autistics, there are nonetheless some intriguing characteristics that many autistics share with computers. Both tend to have difficulty with social-emotional cues. Both tend to be poor at generalizing what they learn, and learn best from having huge numbers of examples, patiently provided.   Both can be fabulous at certain pattern recognition tasks. Autistics, like computers, also may have very good visual memories. Many autistics have indicated that they like interacting with computers, and some have indicated that communicating on the web “levels the playing field” for them, since emotion communication is limited on the web for everyone.

Because many of the issues we face in giving computers skills of emotional intelligence are similar to those faced by therapists working with autistics, we have begun collaboration with these experts. Current intervention techniques for autistic children suggest that many of them can make progress recognizing and understanding the emotional expressions of people if given lots of examples to learn from and extensive training with these examples.

We have developed a system—“ASQ: Affective Social Quotient”—aimed at helping young autistic children learn to associate emotions with expressions and with situations.   The system plays videos of both natural and animated situations giving rise to emotions, and the child interacts with the system by picking up one or more stuffed dwarfs that represent the set of emotions under study, and that wirelessly communicate with the computer. This effort, led by my student Kathi Blocher, has been tested with autistic kids aged 3-7. Within the computer environment, several kids showed an improvement in their ability to recognize emotion [1]. More extensive evaluation is needed in natural environments, but there are already encouraging signs that some of the training is carrying over, such as reports by parents that the kids asked more about emotions at home, and pointed out emotions in their interactions with others. Despite these successes, this work is only one small step; the difficulties in teaching an autistic to appropriately respond to an emotional situation are vast, and we will no doubt face similar difficulties for a long time in trying to teach computers how to respond appropriately.

5 Respectfully handling emotions

Not only do many people feel frustration and distress with technology, but also they show it. A widely publicized 1999 study by Concord Communications in the U.S. found that 84% of help-desk managers surveyed said that users admitted to engaging in “violent and abusive” behavior toward computers. A survey by Mori of 1250 people who work with computers in the UK reported that four out of five of them have seen colleagues hurling abuse at their PC’s, while a quarter of users under age 25 admitted to having kicked their computer. It seems that no matter how hard researchers work on perfecting the machine and interface design, frustration still occurs. In fact, even if computers were as smart as people, they would still sometimes frustrate people; the same is true in human-human interaction: even the most intelligent people sometimes frustrate others. Hence, there is a need to address frustration at run-time – detecting it, and responding to it.

This need is particularly important in light of the impact of stress on health, and the important role of computers in increasing interacting with patients. In some cases, patients prefer giving information to a computer instead of to a doctor, even when they know the doctor will see the information: computers can go more slowly if the patient wishes, asking questions at the patient’s individual speed, not rushing, not appearing arrogant, offering reassurance and information, while allowing the physician more time to focus on other aspects of human interaction [3] [2]. Also, in some cases, patients have reported more accurate information to computers; those referred for assessment of alcohol-related illnesses admitted to a 42% higher consumption of alcohol when interviewed by computer than when interviewed for the same information by psychiatrists [15].

Suppose that a computer could detect patient stress or frustration with high confidence, or that a person directly reports frustration to the machine so that some kind of response by the machine might be appropriate. How should the computer respond? Based on theory that human-computer interaction largely follows the rules of human-human interaction [20], it is germane to explore how a successful human would respond and see if we can find a machine-appropriate way to achieve a similar effect. “It looks like things didn’t go very well,” and “We apologize to you for this inconvenience” are example statements that people use in helping one another manage frustration once it has occurred. Such kinds of statements are known to help alleviate strong negative emotions such as frustration or rage. But can a computer, which doesn’t have feelings of caring, use such techniques effectively to help a user who is having a hard time? To investigate, we built an agent that practices some skills of active listening, empathy, and sympathy, according to the following strategy:

Goal:
Reduce user frustration once it has occurred

Strategy:

1. Recognize (with high probability) that the situation may be frustrating, or that the user is showing signs of frustration likely due to the system
2. Is user willing to talk? If so, then

Practice **active listening,** with **empathy** and **sympathy, e.g.,**
“Good to hear it wasn’t terribly frustrating”
“Sorry to hear your experience wasn’t better”
“It sounds like you felt fairly frustrated playing this game.
Is that about right?”
Allow for repair, in case computer has “misunderstood”
In extreme cases, the computer may even apologize:
“This computer apologizes to you for its part in ..”

3. Polite social closure

Figure 11: The 2x3 experimental design, comparing the emotion-support agent and two control conditions in both a low-frustration condition and a high-frustration condition.

In developing this system, we avoided language where the computer might refer to itself as "I" or otherwise give any misleading implications of having a "self." The system assesses frustration and interacts with the user through a text dialogue box (with no face, voice, fancy animation or other devices that might provoke anthropomorphism.) The only aspect of the interaction that evokes another person is the use of language, which although cleansed of references to self, nonetheless was made deliberately friendly in tone across all control and test conditions, so that friendliness would not be a factor in this study.

The emotion support agent was tested with 70 users who experienced various levels of frustration upon interacting with a simulated network game [14]. We wanted to measure a strong behavioral indication of frustration, since self-report is notoriously unreliable. Thus we constructed a situation where people were encouraged to do their best while test-playing an easy and boring game, both to show their intelligence, and to win one of two monetary prizes. Half of the subjects were exposed to an especially frustrating situation while they played (simulated network delays, which caused the game to freeze, thereby thwarting their attempt to show their intelligence or win a prize). Afterward, subjects would interact with the agent, which was designed to help them reduce their frustration. Finally, they would have to return to the source of their frustration and engage again with the game, at which point we measured how long they continued to interact with it. Our prediction was based on human-human interaction: if somebody frustrates you, and you are still highly frustrated when you have to go back and interact with them, then you will minimize that interaction; however, if you are no longer feeling frustrated, then you are likely to interact with them longer. The 2x3 experimental design is shown in Figure 6, where thirty-four users played the game in a low-frustration condition, while thirty-six played the same game with simulated delays.

We ran three cases for each of the low and high frustration conditions. The first two cases were controls, text-based friendly interactions having essentially the same length as the emotion-support agent. The first control (*ignore*) just asked about the game, ignoring emotions, and the second control (*vent*) asked about the game, but then asked questions about the person's emotional state and gave them room to vent, with no active listening, empathy or sympathy. After interacting with one of the three (*ignore*, *vent*, or *emotion-support*), each player was required to return to the game, and to play for three minutes, after which the quit button appeared and they could quit or play up to 20 minutes more. Compared to people in the *ignore* and *vent* control groups, subjects who interacted with the *emotion-support* agent played significantly longer, behavior indicative of a decrease in frustration.   People in the *ignore* and *vent* cases both left quickly, and there was no significant difference between their times of play. We also analyzed the data to see if there were any significant effects with respect to gender, trait arousability, and prior game playing experience; none of these factors were significant. (For more details regarding this system, experiment, and findings, see Klein (1998) [13]).

These results suggest that today's machines can begin to help reduce frustration, even when they are not yet smart enough to identify or fix the cause of the frustration. Our findings further indicate that it takes very little time to help the user reduce stress – the emotion savvy agent took no more time than the two controls, and all of the interactions took around 4-6 minutes. This time included not only addressing the person's feelings, but also asking several questions about the game. In other words, less than a few minutes of addressing the emotion were sufficient to provide a significant behavioral change in the user.

Today, physicians usually have so little time with patients that they feel it is impossible to build rapport and communicate about anything except the most obviously significant medical issues. However, given findings such as those highlighted at the start of this paper, emotional factors such as stress, anxiety, depression, and anger can be highly significant medical factors, even when the patient might not mention them. Although our findings of a computer's ability to reduce stress were only based on one kind of stress-provoking situation, the strong behavioral effect we obtained in just a few minutes of addressing the emotion, suggests that perhaps something significant can be done by physicians to address emotions related to health, even within the limits of a brief office encounter. If a computer in a few minutes can produce a significant behavioral effect, how much more effect could a truly sensitive and caring emotion-savvy person have in the same amount of time?

6 Concluding remarks

This paper has highlighted several research projects in the MIT Media Lab's Affective Computing Research group. The selected projects are believed to be relevant for future health because they advance the state of the art in physiological sensing, in recognition of emotional signals, in development of emotional skills, and in use of computers to help people manage emotions. Based on the growing number of studies showing that emotions such as anger, anxiety, depression and stress are significant medical factors, helping people better manage these emotions becomes a key form of preventive medicine. As computers assist in gathering information from patients, in helping medical patients communicate with one another and with care-providers, and in disseminating information to patients, the need grows for affective intelligence in the computer interface. Patients who have their feelings properly addressed are more likely to leave satisfied, are less likely to return as often, and are less likely to incur legal cost [7]. These effects translate into dollars saved, so that respecting and responding to patients' emotions is good medicine and good business.

Research into the development of affective technologies is relatively new, and many other labs have recently started similar projects, so that it would take a much longer paper to overview all the research in this area. There are also many exciting findings relating affect and cognition, such as those of Isen and colleagues showing a significant impact of a mild positive state on medical decision making – facilitating efficiency and thoroughness in medical diagnostic reasoning [4], with a number of other benefits [12]. Readers who are interested in related work are encouraged to visit the references of the papers cited at the end of this document, which contain over a hundred pointers to related research conducted beyond our lab.

Over the years, scientists have aimed to make machines and technologies that are intelligent and that help people be intelligent. However, they have almost completely neglected the role of emotion in intelligent interaction, leading to an imbalance where emotions are typically ignored. Similarly, emotions have been largely ignored in the general medical community, with the exception of many recent investigations that have measured their impact and found emotion to be a significant factor in health recovery and in disease prevention. I do not wish to see the scale tilted the other way, where machines twitch at every emotional expression, or where physicians treat emotion and not the accompanying medical problems. What is needed is a reasonable balance. The aim of new affective technologies for medicine should be to help medical care givers attend to patients' full health needs – both emotional and non-emotional–in a balanced, respectful and intelligent way.

7 References

[1] Blocher, K., "Affective Social Quotient (ASQ): Teaching Emotion Recognition with Interactive Media & Wireless Expressive Toys," MIT SM Thesis, May 1999.

[2] Card, W. I. and Lucas, R. W., "Computer interrogation in medical practice," *Int. J. Man-Machine Studies*,**14,** 49-57, 1981.

[3] Dove, G. A. W., Wigg, P., Clarke, J.H.C., Constantinidou, Maria, Royappa, B. A., Evans, C. R., Milne, J., Goss, C., Gordon, M. and de Wardener, H. E., "The therapeutic effect of taking a patient's history by computer," *J. of the Royal College of General Practitioners*, **27**, 477-481, 1977.

[4] Estrada, C. A., Isen, A. M., and Young, M. J. "Positive affect facilitates integration of information and decreases anchoring in reasoning among physicians." *Organizational Behavioral and Human Decision Processes***, 72**, 117-135, 1997.

[5] Fernandez, R. and Picard, R. W., "Signal Processing for Recognition of Human Frustration," *Proc. IEEE ICASSP '98*, Seattle, WA, 1997.

[6] Fernandez, R. and Picard, R. W., "Modeling Drivers' Speech under Stress," *Proc. ISCA Workshop on Speech and Emotions*, Belfast, Sep 2000.

[7] Goleman, D. , *Emotional Intelligence*, Bantam Books, New York, 1995.

[8] Healey, J., Dabek, F. and Picard, R. W., A New Affect-Perceiving Interface and its Application to Personalized Music Selection, *Proc. 1998 Workshop on Perceptual User Interfaces*, San Francisco, CA, 1998.

[9] Healey, J. and Picard, R.W. , "Digital Processing of Affective Signals," *Proc. Int. Conf. on Acoustics, Speech, and Signal Processing*, Seattle, WA, May 1998.

[10] Healey, J. and Picard, R. W. , StartleCam: A Cybernetic Wearable Camera, *Proc. Intl. Symp. on Wearable Computing*, Pittsburgh, PA, 1998.

[11] Healey, J. A. "Wearable and Automotive Systems for Affect Recognition from Physiology," *Ph.D. Thesis, Dept. of Electrical Engineering and Computer Science*, May 2000, MIT.

[12] Isen, A. M., Positive Affect and Decision Making, in Handbook of Emotions, Eds. M. Lewis and J. Haviland, Guilford Press, NY, 2000.

[13] Klein, J., "Computer Response to User Frustration," MIT SM Thesis, Media Arts and Sciences, June 1998.

[14] Klein, J., Moon, Y, and Picard, R. W., "This Computer Responds to User Frustration." *CHI 99 Short Papers*, Pittsburgh, PA. A longer version has been accepted for publication in *Interaction with Computers*.

[15] Lucas, R. W., Mullin, P. J., Luna, C. B. X., and McInroy, D. C., "Psychiatrists and a Computer as Interrogators of Patients with Alcohol-Related Illnesses: A Comparison," *Brit. J. Psychiat.*, **131**, 160-7, 1977.

[16] Marrin, T. and Picard, R. W. , "Analysis of Affective Musical Expression with the Conductor's Jacket," *Proc XII Col. Musical Informatics*, Gorizia, Italy, 1998.

[17] Marrin, T., "Inside the Conductor's Jacket: Analysis, Interpretation, and Musical Synthesis of Expressive Gesture," MIT PhD. Thesis, Media Arts and Sciences, 1999.

[18] Picard, R. W., *Affective Computing*, MIT Press, Cambridge, MA, 1997.

[19] Picard, R. W, and Healey, J., Affective Wearables, Personal Technologies **1**, No. 4 , 231-240, 1997.

[20] Reeves B. and Nass, C. (1996). The Media Equation, Cambridge University Press; Center for the Study of Language and Information.

[21] Riseberg, J., Klein, J., Fernandez, R. and Picard, R.W., "Frustrating the User on Purpose: Using Biosignals in a Pilot Study to Detect the User's Emotional State," *CHI '98 Short Papers*, Los Angeles, CA. A longer version of this paper has been accepted for publication in *Interaction with Computers,* 1998.

[22] Scheirer, J., Fernandez, R. and Picard, R.W., "Expression Glasses: A Wearable Device for Facial Expression Recognition," *CHI '99 Short Papers,* Pittsburgh, PA, 1999.

[23] Sheridan, T. B., "Community Dialog Technology", *Proc. of the IEEE*, 63, No. 3, March 1975.

[24] Vyzas, E. "Recognition of Emotional and Cognitive States Using Physiological Data," *Mechanical Engineer Thesis,* June 1999, MIT.

[25] Vyzas, E., and Picard, R. W. , "Online and Offline Recognition of Emotion Expression from Physiological Data," *Workshop on Emotion-Based Agent Architectures at the Int. Conf. on Autonomous Agents*, Seattle, WA, 1999.