

Toward computers that recognize and respond to user emotion

by R. W. Picard

For a long time emotions have been kept out of the deliberate tools of science; scientists have expressed emotion, but no tools could sense and respond to their affective information. This paper highlights research at the MIT Media Laboratory aimed at giving computers the ability to comfortably sense, recognize, and respond to the human communication of emotion, especially affective states such as frustration, confusion, interest, distress, anger, and joy. Two main themes of sensing—self-report and concurrent expression—are described, together with examples of systems that give users new ways to communicate emotion to computers and, through computers, to other people. In addition to building systems that try to elicit and detect frustration, our research group has built a system that responds to user frustration in a way that appears to help alleviate it. This paper highlights applications of this research to interface design, wearable computing, entertainment, and education, and briefly presents some potential ethical concerns and how they might be addressed.

Not all computers need to “pay attention” to emotions or to have the capability to emulate emotion. Some machines are useful as rigid tools, and it is fine to keep them that way. However, there are situations in which human-computer interaction could be improved by having the computer adapt to the user, and in which communication about when, where, how, and how important it is to adapt involves the use of emotional information.

Findings of Reeves and Nass at Stanford University suggest that the interaction between human and machine is largely natural and social,¹ indicating that

factors important in *human-human* interaction are also important in *human-computer* interaction. In human-human interaction, it has recently been argued that skills of so-called “emotional intelligence” are more important than are traditional mathematical and verbal skills of intelligence.² These skills include the ability to recognize the emotions of another and to respond appropriately to these emotions. Whether or not these particular skills are *more* important than certain other skills will depend on the situation and goals of the user, but what is clear is that these skills are important in human-human interaction, and when they are missing, interaction is more likely to be perceived as frustrating and not very intelligent.

Consider an example of human-human interaction:

Suppose that a human starts to give you help at a bad time. You try ignoring, then frowning at, and then maybe glaring at him or her. The savvy human infers you do not like what just happened, ceases the interruption, notes the context, and learns from the feedback.

This scenario is easily modified for human-computer interaction and is relevant to the growing use of software that is designed to interrupt a person with reminders, hints, and offers of suggestions. Now we take out the word “human” in the example and replace it with the word “computer”:

©Copyright 2000 by International Business Machines Corporation. Copying in printed form for private use is permitted without payment of royalty provided that (1) each reproduction is done without alteration and (2) the *Journal* reference and IBM copyright notice are included on the first page. The title and abstract, but no other portions, of this paper may be copied or distributed royalty free without further permission by computer-based and other information-service systems. Permission to *republish* any other portion of this paper must be obtained from the Editor.

Suppose that a computer starts to give you help at a bad time. You try ignoring, then frowning at, and then maybe glaring at it. The savvy computer infers you do not like what just happened, ceases the interruption, notes the context, and learns from the feedback.

Of course, the “savvy computer” does not exist yet. Today’s computers largely ignore the user’s signs of frustration. This situation is similar to having a person ignore your signs of growing irritation, which would likely lead you to think of that person as annoying and rather stupid—unless, perhaps, you were highly intimidated by the person, in which case you might blame your own actions. An analogy with human-computer interaction also holds, where people think of computers as annoying and stupid, or as intimidating, and of themselves as dummies. The similarity in the interaction holds despite the fact that people know that computers are not humans and that computers do not have the same capabilities as humans. We all know better than to treat computers like humans, and yet many of our default behaviors and responses tend in this direction.

Giving computers skills of emotional intelligence in the broad sense originally described by Salovey and Mayer³ involves more than giving them the ability to recognize and respond to human emotions. It involves other aspects of affective computing—computing that relates to, arises from, and deliberately influences emotion—as well as other nonaffective capabilities, such as sensing of and reasoning about a context. A fairly extensive overview of research related to affective computing, up to and including early 1997, can be found in the author’s book.⁴ This paper does not attempt to perform an overview of the field since publication of the book but rather focuses on an overview of affect *communication* research at the Massachusetts Institute of Technology (MIT) Media Laboratory since 1997. The reader who is interested in learning more about the latest affect-related work at other locations might turn to the numerous recent workshops and special sessions that address a broad variety of aspects of emotion and computation related to artificial intelligence, software agents, and human-computer interfaces; several of these have published or soon-to-be published proceedings.^{5–10} Also, the references given at the end of this paper contain many citations to related work done outside the MIT Media Lab.

Although the work of this paper focuses on recognition and response to user emotion, there are other

aspects of affective computing, such as the use of emotion-like mechanisms within the machine to help make decisions, adjust perception, and influence learning. Even though the author is also involved in some of that research going on outside the Media Lab at MIT,^{4,11,12} discussion of those areas of affective computing will not be included in this paper. Here, the focus is on the area of giving machines new abilities to recognize and respond to emotion, without actually giving them emotion-like mechanisms.

In the Affective Computing Research Group at MIT, we are particularly interested in the intelligent handling of affective states commonly expressed around computer systems: frustration, confusion, dislike, like, interest, boredom, fear, distress, and joy. The idea is that if the computer detects expressions of these states, and associates them with what it is doing and with other events in the environment, then it can better learn which interactions are successful and which need improvement. In other words, the computer can directly collect usability information related to a user’s emotion, such as what parts of the interface or system were in use when the user expressed the greatest frustration. This information can be passed on to a designer, who can try to use it to improve future versions of the system. Alternatively, if the system itself is “smart” and entrusted with its own adaptation, then it can use the affective feedback to try to improve its behavior. The latter is riskier, in the sense that few systems are capable yet of doing this in a way that would not simply cause the user’s frustration to escalate. Successful adaptation to users is an important research area for machine learning and human-computer interaction, and I expect to see future systems move toward the latter aim as well as the former.

The rest of this paper is divided into three subject areas: (1) enabling the user to communicate emotion in a way that is physically and psychologically comfortable, generally by offering several possible styles of direct and indirect sensing and new tools for communication; (2) reducing user frustration, with a focus on how the computer can handle user frustration once it has occurred; and (3) developing applications that handle affective information, with examples from several domains.

Comfortable communication of emotion

Emotion communication requires that a message be both sent and received. Most computer interfaces inhibit such communication. Several people with au-

tism, a complex disorder that typically includes impairment in recognition of emotion, have commented that they love being on the Web because it levels the playing field for them. In a sense, everyone is autistic on line. With the exception of gifted poets and others who work hard to lessen the ambiguity of the emotions expressed by their text e-mails, most of the emotions we show to our keyboards, monitors, and mice are not transmitted. Sometimes this is good; however, often it is a source of miscommunication and misunderstanding, resulting in lost time, damaged relationships, and reduced productivity.

Our research group is building tools to facilitate multiple kinds of emotion expression by people, not to force this on anyone, but to allow for a larger space of possibilities for those who want to communicate affective information. The tools include new hardware and software that we have developed to enable certain machines not only to receive emotional expression, but also to recognize meaningful patterns of emotional expression.

Emotion communication can be conducted with varying degrees of naturalness and accuracy, and with some surprises about what may be considered most preferable. I do not advocate any one means of communication, but believe in providing an array of choices so that users who wish to communicate emotion can pick what they prefer. Two main themes of communication run through the kinds of devices we have built at the Media Lab: *self-report* and *concurrent expression*. These two themes are described next, together with an example of a human-human analogy, some pros and cons, and examples of new systems that our group has designed and built.

Self-report. Self-report systems leave it up to the user to go out of his or her way to communicate emotion. The user might make a selection by means of software from a menu of emotions with words or icons, or the user might touch a hardware input device that acts as a tangible icon, e.g., a physical icon, or “phicon” in the language of Ishii and Ullmer.¹³ In either case, it is up to the user to select an item when he or she feels a certain way and wants the system to know. (If natural language processing were sufficiently advanced, the system could accept freely typed or spoken input instead.) Following are characteristics of self-report systems:

- A human-human analogy of these systems: One person interrupts a conversation or task to clearly state how he or she is feeling.

- The pros of these systems: This option gives the user total conscious control over the message that is sent, and sometimes it is a natural way to express emotion.
- The cons of these systems: The user has to stop the primary task he or she is doing in order to communicate emotion. It is often hard or burdensome to articulate one’s feelings. The menu of words, icons, or phicons can become large and still not capture what one is feeling.

An example of self-report—thumbs-up feedback. To provide ongoing usability feedback, Matt Norwood in the MIT Media Lab has built a thumbs-up and a thumbs-down phicon and attached both to the networked “Mr Java” coffee machine in our lab (Figure 1). This machine usually works fine, making good cappuccino and espresso, but it occasionally displays cryptic messages. Norwood has added the ability to associate user feedback with the status of the machine’s functions, including error states, such as “out of beans” and “out of milk.” When a customer is satisfied, he or she can tap the thumbs-up phicon, and the system will record a notch of satisfaction for the current operating state. If the machine displays an error message that the customer does not understand, such as “empty grounds bin,” the customer might tap the thumbs-down phicon. The system records the continuous stream of reports of satisfaction or dissatisfaction, associating each with one of the 31 operating states of the machine at the time of feedback. This information provides the designer of the machine with a usability report of which features pleased or displeased the most customers, based on when users chose to express such information. Details of our findings with this device can be found in Norwood’s thesis.¹⁴

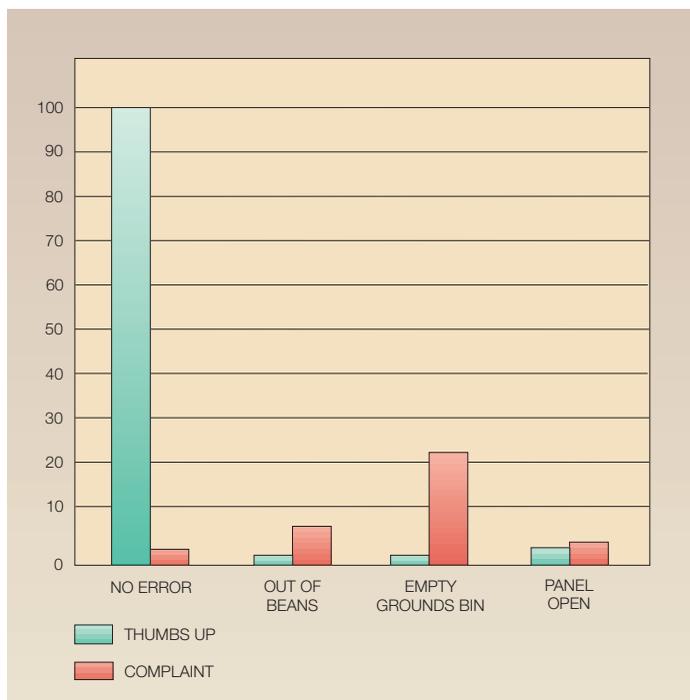
Concurrent expression. In concurrent expression, the system attempts to sense affective expression in parallel with the user’s primary task, without the user having to stop what he or she is doing to report his or her feelings. This can happen via whichever sensors the user may choose for the computer to have: video, microphone, typing or mouse holding pressure, physiology, olfaction, and so forth. Characteristics of concurrent expression follow:

- A human-human analogy: A person expresses emotion in whatever way is most natural for the given situation, while engaged in some task, without having to stop that task.
- The pros: This method aims at the greatest naturalness, placing no additional burden on the user.

Figure 1 At left: the customer taps the thumbs-up icon after getting a frothy cup of cappuccino; at right: a bar graph illustrating positive feedback (green) and negative feedback (red) received as a function of system operating state



Photo by Webb Chappell, copyright 1999 Webb Chappell; reprinted with permission



The user does not have to be interrupted, nor does he or she have to put into words something that might be difficult to express verbally.

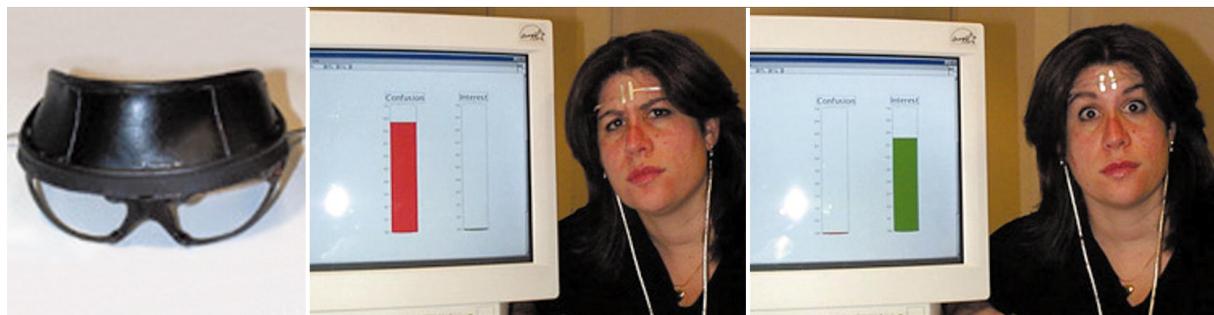
- The cons: Users may feel uneasy about a machine sensing things that they themselves are unaware of having expressed. There is more room for misinterpretation, given that much nonverbal information is ambiguously communicated. Many forms of sensing may be considered obtrusive, especially from a privacy standpoint.

One example of a concurrent-expression method—expression glasses. In many cases, such as a human factors assessment of users engaging a complex product, a direct means of assessment is fine *after* the interaction, but not during it, when it could interfere with the person's task. Consider a focus group in which participants are asked to indicate the clarity of packaging labels. If while reading line three, a subject furrows his brow in confusion, he has communicated concurrently with the task at hand. Brow furrowing can be detected by a camera if lighting and head position are 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 is a pair of wearable “expression glasses” (Figure 2) that sense changes in facial muscles, such as furrowing the brow in confusion or interest.¹⁵ These glasses have a small point of contact with the brow, but otherwise can be considered 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 concurrently, *while concentrating* on a task, and can be activated either unconsciously or consciously. People are still 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 (which I think is good), but if they want to communicate, the glasses offer a virtually effortless way to do so. Details on the design of the glasses and on experiments conducted with the expression glasses are available in the paper by Scheirer et al.¹⁵

Figure 2 Expression glasses (left) shown with “stealth” visor for privacy; Jocelyn Scheirer (center and right) wearing sensors that detect an expression of confusion, center (red), and of interest, right (green)



Note that the two main forms of communication just described, self-report and concurrent expression, can be used together. For example, the standard self-report means of giving a subject a questionnaire might be useful after a scenario like the one just described. The two methods complement each other in that they can gather different information. For example, self-report is notoriously inaccurate for getting true feelings (affected by a subject's expectations, mood, comfort level, rapport with the experimenter, and so forth) but can still provide useful feedback by helping articulate causes of feelings and other important issues.

Why wear expression glasses, instead of raising your hand or pushing a button to say you are interested or confused? The answer is not that there should be one *or* the other; both have a distinct role. Sometimes the subtlety of this point is missed: affect is continuously communicated while you are doing just about anything. When you pick up a pen, you will do so very differently when you are angry than when you are joyful. When you watch somebody, your eyes will 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 with these concurrent expressions, if you want to push a button to communicate your feelings, or if you want to raise your hand to say you are confused, you have to interrupt what you were doing beforehand.

In studies conducted decades ago, participants were given self-report buttons that they could use in a classroom or meeting to anonymously communicate their feelings and opinions. The users of these sys-

tems describe the technology as helping to engage them more in the interaction, believing that their opinions mattered more.¹⁶ This method was successful when used in a question-and-answer format, e.g., “Did you all like this?” with everyone then pressing “yes” or “no.” However, when the device was used to communicate feedback such as confusion to a professor in a classroom, there were several problems. The key problem was that while a person was concentrating on what was confusing and trying to understand it, he or she would forget to push the button. Self-report is useful when there is a break in the action, with time to ask about and assess what has happened. But at that point, it takes more work to relate the expressed confusion to precisely where it was triggered. In contrast, frowning the brow happens without necessarily interrupting the flow of action; the listener can change his or her facial expression without having to think about doing so; a person can concentrate on the lecture instead. Self-report is important, but it is no substitute for the natural channels of largely nonverbal communication that humans use concurrently while engaged in conversation, learning, and many other activities.

Another example of a concurrent-expression method—physiological analysis. Our lab has been active in developing new interfaces for future computing, including wearable computers that would offer a highly personalized computing environment, which attends not just to a person's direct input but also to the person's behavioral and affective cues.¹⁷ Toward this goal, we have been developing new wearable sensors that attempt to be more comfortable physically and socially (look more like fashion accessories than like clunky medical instrumentation), facilitating the gathering of affective information concurrent with

day-to-day activities.¹⁸ A wearable computer affords a different kind of sensing; in particular, it is relatively easy to obtain signals from the surface of the skin, in contrast with a seated environment where it is easier to point a camera at the user. We have been developing algorithms that can read four physiological signals, comfortably sensed from the skin surface, and relate these to a deliberately expressed

Physiological expression is one of many multimodal means of concurrent emotion communication we are exploring.

emotion. Our recent results have achieved 81 percent recognition accuracy in selecting which of eight emotions was expressed by an actor, 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 Vyzas and Picard¹⁹ for details of the data collection and the algorithms we developed for recognition.) The eight emotions investigated were: neutral, hatred, anger, romantic love, platonic love, joy, and reverence. These results are the best reported 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 our results are only for a single user, and they are obtained by a forced selection of one of the eight categories. Hence, these results are comparable to the 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 (or their affective expression) from different cultures, of different ages, speaking (or expressing) continuously, in a variety of different environments.

Physiological expression is one of many multimodal means of concurrent emotion communication that our group is exploring. We are also beginning to analyze affect in speech, an area in which humans only perform at about 60 percent accuracy (on roughly eight emotion categories, when the content of the speech is obscured). Our initial focus is on speech from automobile drivers who might be conversing on the phone or with an automobile navigation system. We have recently equipped a car to examine driver behavior features jointly with physiological information.²⁰ One such sensor setup is shown in Figure 3 by which the following physiological indicators were measured: muscle tension with an electromyogram (EMG), blood volume pressure (BVP), skin conductivity (SC), and respiration. All these efforts gather data that push the abilities of traditional pattern recognition and signal processing algorithms, which have difficulty handling the day-to-day and interpersonal variations of emotional expression. Consequently, we are also conducting basic research in machine learning theory and in pattern recognition to develop better methods for modeling and inference. I expect that affective systems will use affect as input to a machine learning system, which will not only try to learn to do better at recognizing affect but will also learn to do better at adapting to the user, based on his or her positive or negative feedback.

It is important to keep in mind that some people will feel more or less comfortable with each of these forms of communication. In a survey of nine regular users of traditional computing involving a big software package at MIT, with whom we were working to find ways to gather affective feedback about the product, five of them wanted a self-report button, whereas the other four wanted a beanbag, special surface, or mouse pressure sensor that they could hit, squeeze, or bang on. Such input, if mapped to an icon on the screen that registers disapproval, could be viewed as a nonverbal means of self-report, which the user was in control of at all times. If the mouse were the primary source of such hitting, or squeezing, or banging input, then the feedback could also occur concurrently with the task, giving the user another option. We thus built a mouse that senses pressure patterns but is only activated when the user directly points the mouse at a feedback icon (Figure 4). This mouse is part of new work by Carson Reynolds, who is developing an “Interface Tailor” that would adapt to user displays of positive or negative feedback. The mouse is a natural place for sensing signals from the user’s hand; Dana Kirsch in our lab

Figure 3 Sensor placement in experiments to learn about driver stress and other affective states that might influence driving safety



built an earlier version for trying to sense positive or negative information,²¹ and IBM has built a version for sensing six emotions.²²

It was interesting to discover that in our survey of the nine users, video sensing was not selected by any of them, and only one person selected microphone monitoring of vocal stress. Although facial and vocal expressions are the best-known social forms of human emotional expression, and new forms of sensing are beginning to offer more natural emotion communication, many people are not ready for this type of sensing. Users have strong feelings about if, when, where, and how they want to communicate their emotions, as well as about the privacy of their identifying and appearance information (face and voice). It would be absurd if developers of affective computing, which is ultimately about respecting human feelings, did not respect people's feelings in the deployment of this technology. Our group's ethic insists on giving users choices and control.

Figure 4 Mouse equipped with resistive foam that senses pressure patterns to discern everyday gentle handling from pounding and other actions



Photo by Webb Chappell, copyright 1999 Webb Chappell; reprinted with permission

The above emphasis was on human-computer affect communication. Most of this also applies to human-human communication when mediated by a computer, although in the case of the latter it is not nec-

Figure 5 Use of the TouchPhone, illustrated by Scheirer; transmitted color ranges from blue, through green, yellow, orange, and finally red, as the pressure applied to the phone increases



essarily desirable that the machine interpret or recognize the emotional signals, only that the machine facilitate the priority and accuracy of their transmission. An example of a system designed explicitly to expand human-human communication capabilities via computer is the TouchPhone, developed by Jocelyn Scheirer in our group (Figure 5). The TouchPhone augments regular voice communication with pressure information, indicating how tightly the speaker is holding the phone. The pressure is 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. No interpretation of this signal is performed by the computer; the color signal is simply transmitted to the conversational partner as an additional channel of information.

I have met with four of my students over four hours of TouchPhone conversations, and the results, although anecdotal, were interesting and were consistent with other experiences we have had with this and other emotion-communication technologies we have been developing. I found that each of the four students had a nearly unique color flicker pattern, which was distracting at first. After I moved the color pattern into my periphery, it became an ambient background pattern that was no longer distracting, 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 interpret his feel-

ings from the color, he informed me that he was not trying to squeeze it tighter at all, and he thought it was red because he was stressed about a question I asked him. This person was a very nonexpressive, down-to-earth engineering student 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 his opening up a greater range of emotional communication, by his choice; it did not impose this on him, but it made it easier for him. The color was not an expression of how he or any of the students was truly feeling. However, the system provided a new channel of nonverbal communication that could, in turn, open up a new line of verbal communication. The TouchPhone is a new vehicle for carrying information that may change with emotion, but it will not reveal a person's emotion.

Reducing user frustration

Not only do many people feel frustration and distress with technology, but they also show it. A widely publicized 1999 study by Concord Communications in the United States found that 84 percent 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 U.K. reported that four out of five of them have seen colleagues hurling abuse at their PCs, and 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, frus-

tration can occur in the interaction. Researchers of human-computer interaction have worked hard to prevent frustration, which continues to be an important goal. However, despite their best efforts, an unforeseen situation often creates stress in the interaction. Even if computers were as smart as humans, they would still sometimes behave in a frustrating way, because even the most intelligent humans sometimes behave in a frustrating way. Hence, there is a need to address frustration at run time—detecting it, and responding to it.

I originally thought it would be easy to collect data from frustrated users. One could just ask them to sit in front of a computer running a particular software package, and—*voilà!* Alternatively, one could hire an actor to express emotions, and record them. If the actor used method acting or another technique to try to self-induce true emotional feelings, the results would closely approximate emotions that arise in natural situations. A student using visualization techniques for feeling different emotions was the method we used in collecting 30 days of physiology data.^{19,23} However, these examples are not as straightforward as they may seem at first: they are complicated by issues such as the artificiality of bringing people into laboratory settings, the mood and skill of an actor, whether or not an audience is present, the expectations of subjects who think you are trying to frustrate them, the unreliability of a given stimulus for inducing emotion, the fact that some emotions can be induced simply by a subject's thoughts (over which experimenters have little or no control), and the sheer difficulty of accurately sensing, synchronizing, and understanding the "ground truth" of emotional data. In response to these difficulties, we have begun to develop lab-based experimental methodologies for eliciting and gathering frustration data.

Building a system. In one methodology that we developed, we built a system that attempts to elicit and record signs of frustration from a user in a carefully controlled way using concurrent expression.²⁴ This system involves users in an experiment whereby they try to solve a sequence of easy, visual perception puzzles as accurately and as quickly as possible in order to win a \$100 prize. During the task, we introduce randomly appearing delays, where the mouse seems to stick and the screen does not advance to the next puzzle, thus damaging their score. Such an event is likely, although not guaranteed, to elicit frustration.

We compared features from two physiological signals: blood volume pressure and skin conductivity,

during episodes when the mouse "stuck" vs episodes when all was going smoothly. Training hidden Markov models for each of these two categories, for each user from part of the data, Raul Fernandez built an algorithm that tried to infer which category a person was in, given only the physiological features from unseen data. Initial results were significantly better than a random choice at detecting and recognizing the potential frustration episodes in 21 out of 24 users.²⁵ However, the results were still significantly less than 100 percent, indicating that although this information is helpful, it must be combined with other signals for a more confident decision about the affective state.

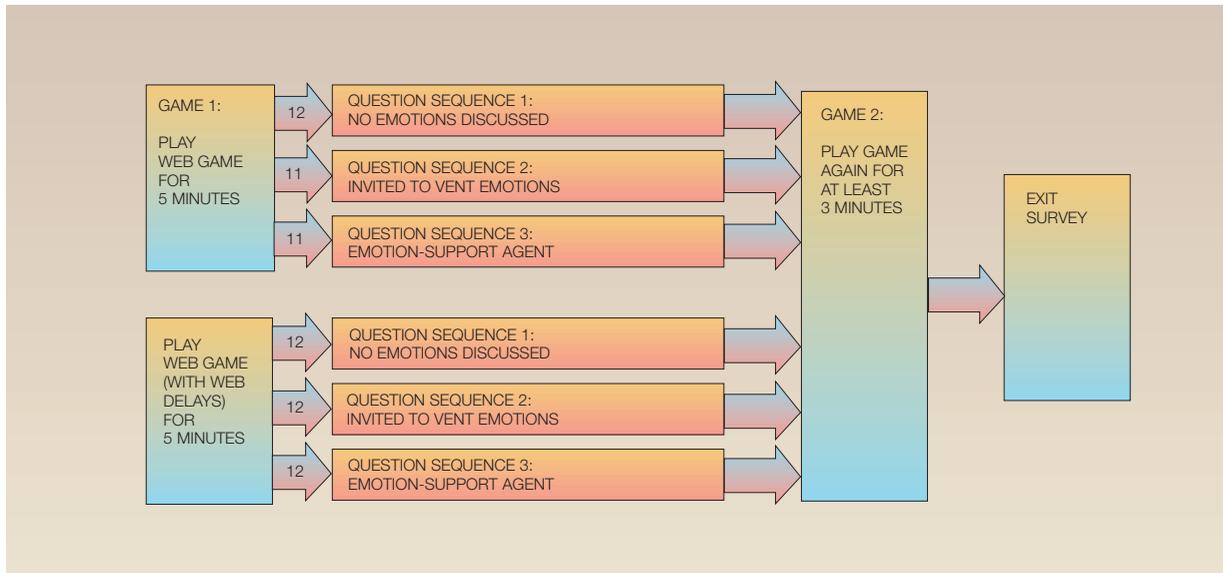
Suppose that a computer could detect frustration with high confidence, or that a person directly reports frustration to the machine so that some kind of response is appropriate. *How* should the computer respond? Returning to the Reeves and Nass arguments, I believe it is important to explore how a successful human would respond, and see whether we can find a machine-appropriate way to imitate this response. "It looks like things did not go very well," and "We apologize to you for this inconvenience" are examples of statements that humans 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 does not 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, which has been designed based on other researchers' analysis of successful human-human interaction:

Research 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 the 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?"

Figure 6 The 2 x 3 experimental design



- Allow for repair, in case the computer has “misunderstood.”
- In extreme cases, the computer may even apologize, e.g., “This computer apologizes to you for its part in . . .”

3. Provide polite social closure.

In developing this system, we took care to avoid language in which 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 dialog box (with no face, voice, fancy animation, or other devices that might provoke anthropomorphism). The only aspect of the interaction that evokes the image of another person is the use of the English 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.

Using the system. The emotion support agent was tested with 70 users who experienced various levels of frustration upon interacting with a simulated network game.²⁶ We wanted to measure a strong behavioral indication of frustration, since self-report is notoriously unreliable. Thus we constructed a sit-

uation where people were encouraged to do their best while test-playing an easy and boring game, both to display 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 display their intelligence or win a prize). Afterward, subjects would interact with the agent, which would attempt 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 someone 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, you are likely to remain with them longer. The 2 × 3 experimental design is shown in Figure 6, where 34 users played the game in a low-frustration condition, and 36 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, both of which were text-based, friendly interactions having essentially the same length as the emotion-support agent: the first (ignore) just asked

about the game, ignoring emotions, and the second (vent) asked about the game, but then asked questions about the person's emotional state and gave him or her room to vent (with no active listening, empathy, or sympathy). After interacting with one of the three cases (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 came on and they were free to quit or to play up to 20 minutes. Compared to people in the ignore and vent control groups, subjects who interacted with the emotion-support agent played significantly longer, demonstrating behavior indicative of a decrease in frustration. People in the ignore and vent cases both left much sooner, and there was no significant difference between their times of play.

These findings held true within the low-frustration condition and within the high-frustration condition: those who interacted with the emotion-support agent played significantly longer than those who interacted with similar agents that ignored their emotions (ignore) or just asked them about their emotion (vent). We also analyzed the data to see whether there were any significant effects with respect to gender, trait arousability, and prior game playing experience; none of these factors was significant. (For more details regarding this system and its human experiment and findings, see Klein et al.²⁶) 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.

These findings were of no surprise to one colleague, visiting from a large computer firm. He described how his company had conducted a study of their customers to find which were most likely to buy their brand of computer again: those who had problems with the computer and received good support, or those who had no problems with the computer. The results showed that those who had problems and received good support were significantly more likely to buy this company's brand again than were customers who had *no* problems with their computers. Their findings underscore the need for good customer support, which usually implies teams that have been carefully trained in the importance of practicing active listening, and in the appropriate use of empathy and sympathy. Our findings show that these skills, when practiced by a computer, can also have a strong behavioral effect.

Our findings in this area, and those findings of the large computer company, are disturbing, not only be-

cause of the increased tendency that companies have to rush products to customers that are barely 80 percent debugged, but because of the potential for deliberate manipulation of customers' feelings. Could computers manipulate customers' feelings? Yes. But would the technique used by the agent succeed repeatedly? Probably not. The rationale lies again in examining the nearest human-human equivalent. If a customer service agent repeatedly showed active listening, empathy, and sympathy for the source of the frustration, but never did anything to help solve the problem, the customer would wise up quickly. In contrast, if the agent played the role of a friend who supports the customer, but does not solve the customer's problems, then that support might still be effective. I do not see the virtual friend technology as ready for commercial deployment anytime soon, and it raises many other concerns, but it is a future possibility, and affect communication would play an important role in its development.

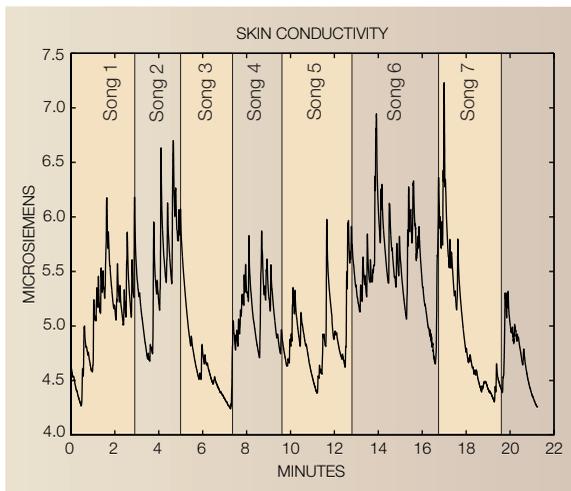
This technology raises potential ethical concerns, most of which are not new to this technology. For example, advertising and persuasive technologies (see special issue of *Communications of the ACM*, May 1999) already focus on manipulating emotion, and many other forms of surveillance and sensing technology, coupled with networked databases, have researchers examining privacy implications. I expect that this technology will succeed best when it is applied only in the service of users' needs, under their control, and for their benefit. The interested reader is referred to the discussions in Klein's thesis²⁷ and in Chapter 4 of the author's book⁴ for specific issues related to sensitive handling of affective information.

Developing applications

Affective computing suggests many new applications and variations on existing applications, in entertainment, in the arts, in education, and in regular human-computer interaction. Many such possibilities have been envisaged in the author's 1997 book (Chapters 3 and 8⁴), and this section will highlight some of the progress we have made since 1997. In each case, it is important to keep in mind that what we are aiming for is not a one-size-fits-all system, but a system that can adapt to each user in a way that is sensitive and respectful of his or her expressed feelings.

As mentioned above, we have designed and built affective wearable computers that sense information from a wearer going about daily activities.¹⁸ Some of these wearables have been adapted to control de-

Figure 7 Skin conductivity for a person listening to seven songs; signal is used by the “Affective DJ” to aid in song selection



VICES for the user, such as a camera that saves video based on one’s arousal response,²⁸ tagging the data not just with the usual time stamp, but also with information about whether or not it was exciting, as indicated by patterns in the user’s skin conductivity. Using similar information, our group has built a wearable “DJ” that not only tries to select music from performers that are preferred, but also attempts to adjust its selection based on a feature of the user’s mood.²⁹ From prior listening, it can watch which songs the user prefers to listen to when calm or active, and which tend to calm or pep up the user more. Figure 7 shows the skin conductivity for one wearer while listening to seven songs and otherwise relaxing. The level is noticeably higher for songs 2 and 6, agreeing with the subject’s report that these songs were perceived as energizing. The level is noticeably lower for songs 3 and 7, agreeing with the subject’s report that these were relatively calming. If the subject chooses, he or she can tell the system such things as “choose some jazz to pep me up,” and it will select music from preferences that have these features.

One thing to be clear about is that this system does not impose choices on the user; for example, it will not try to “pep up” the user unless he or she has instructed it (directly or indirectly) to do so. It remains under the user’s ultimate control, which I consider to be an important factor. The user can turn it on or off, and the user retains control over the types of

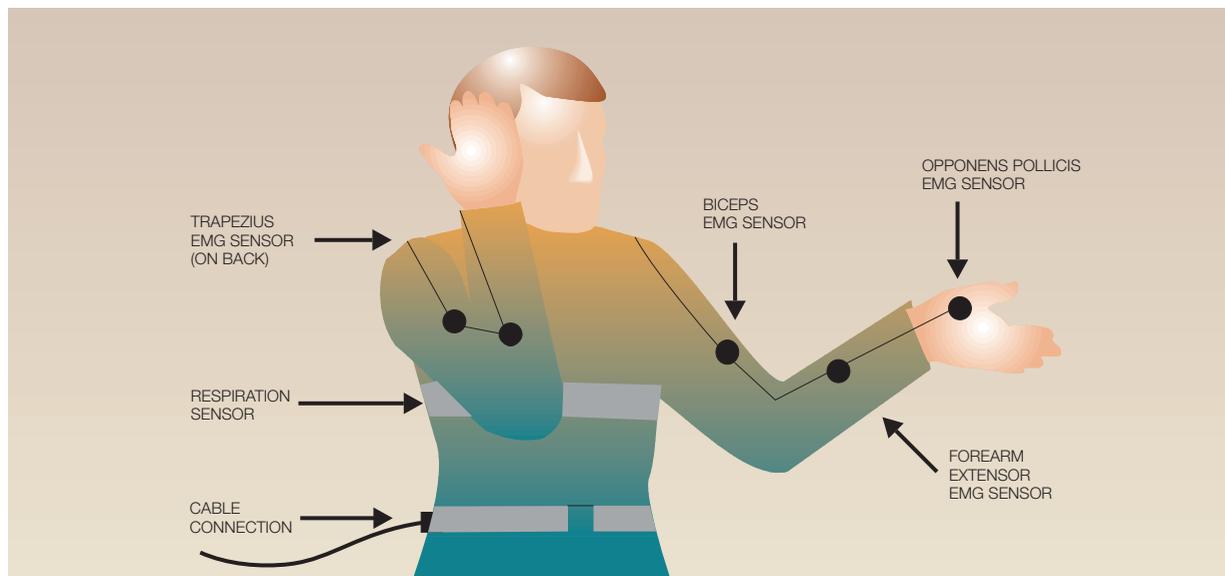
music it will play. Ideally, affective systems should never impose something on their user unless the user indicates that he or she wants the imposition. If affective technology annoys the user, it has probably not succeeded.

An application in the creative arts is the development of a highly expressive wearable system: the Conductor’s Jacket. Figure 8 shows one of the early prototypes designed and built by my research assistant, Teresa Marrin. The jacket maps patterns of muscle tension and breathing into features that shape the music. Seven electromyogram (EMG) and one respiration sensor are included in the version shown here. This wearable system was built 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 about 30 significant expressive features and hypothesized several principles of communication via musical gestures.³⁰ She has subsequently developed a version of the jacket that can transform natural expressive gestures of the wearer into real-time expressive shaping of MIDI (Musical Instrument Digital Interface) performances.³¹ A professional conductor herself, Marrin has performed in the jacket in several large venues with great acclaim. She is currently looking at how the jacket can be used to help educate student conductors, giving them more precise feedback on their timing, tension, and other important aspects of expressive technique.

During the course of this research, we have realized that there are many parallels between autistics, who tend to have severely impaired social-emotional skills, and computers, which do not “get it” when it comes to emotion. Both also tend to have difficulty generalizing, even though both can be fabulous at certain pattern recognition tasks. 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 some 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 that is aimed at helping young autistic children learn to associate emotions with expressions and with situations. The system plays videos of both natural and animated

Figure 8 Schematic of sensors in the Conductor's Jacket



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 Kathi Blocher, has been tested with autistic children ages 3–7 years. Within the computer environment, several children showed an improvement in their ability to recognize emotion.³² More extensive evaluation is needed in natural environments, but there are already encouraging signs that some of the training is carrying over. Nonetheless, this work is only one small step up a huge mountain; 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.

Affective computing is in its infancy, but its potential applications are vast. Our research group is continuing to develop this technology while exploring new application areas: how devices that help communicate emotion can be used in focus groups and in audience-performer interaction, how affect-sensing in the car might improve safety, how wearable devices can boost emotional awareness and help wearers regulate stress and potentially improve health, how emotion can be sensed and responded to respectfully in e-commerce, and how emotion

communication may potentially improve a learning experience for children or adults. Emotion is a natural part of many of life's experiences, but so far computers have largely ignored it. As computers become ubiquitous, and as users demand more customized, intelligent, adaptive interactions, I think it will become increasingly important for affective communication to be enabled in this interaction.

Concluding remarks

This paper has highlighted several research projects in the Affective Computing Research Group of the MIT Media Lab. The emphasis has been on illustrating new technology that can begin to recognize and help communicate aspects of emotional expression and to respond to it in an appropriate way. This area of research is very new, and many other laboratories have recently started similar projects, so that it would take a much longer paper to provide an overview of all the research in this area. Readers who are interested in related work are encouraged to peruse the references of the papers and the workshop proceedings listed at the end of this paper, all of which contain many pointers to related research conducted beyond MIT.

Over the years, scientists have aimed to make machines that are intelligent and that help people use

their native intelligence. However, they have almost completely neglected the role of emotion in intelligence, leading to an imbalance on a scale where emotions are almost always ignored. I do not wish to see the scale tilted out of balance in the other direction, where machines twitch at every emotional expression or become overly emotional and utterly intolerable. However, research is needed to learn about how affect can be used in a balanced, respectful, and intelligent way; this should be the aim of affective computing as we develop new technologies that recognize and respond appropriately to human emotions.

Cited references

1. B. Reeves and C. Nass, *The Media Equation*, Center for the Study of Language and Information, Stanford University (1996).
2. D. Goleman, *Emotional Intelligence*, Bantam Books, New York (1995).
3. P. Salovey and J. D. Mayer, "Emotional Intelligence," *Imagination, Cognition, and Personality* 9, No. 3, 185–211 (1990).
4. R. W. Picard, *Affective Computing*, MIT Press, Cambridge, MA (1997).
5. "Workshop on Grounding Emotions in Adaptive Systems," D. Canamero, C. Numaoka, and P. Petta, Editors, *5th International Conference of the Society for Adaptive Behavior (SAB '98)*, Zurich (August 1998); <http://www.ai.univie.ac.at/~paolo/conf/sab98/>.
6. "Emotional and Intelligent: The Tangled Knot of Cognition," D. Canamero, Editor, *AAAI Fall Symposium*, Orlando, AAAI Press, Menlo Park, CA (October 1998).
7. Workshop on Emotion-Based Agent Architectures (EBAA '99), J. D. Velasquez, Editor, in conjunction with the Autonomous Agents Conference, Seattle (May 1999).
8. *Emotions in Humans and Artifacts*, R. Trappl and P. Petta, Editors, collected papers from Vienna workshop (August 1999); to be published by MIT Press, Cambridge, MA.
9. "Human-Computer Interaction: Ergonomics and User Interfaces," Volume 1, H.-J. Bullinger and J. Ziegler, Editors, special session on Affective Computing and HCI, *Proceedings of HCI International '99*, Munich, Lawrence Erlbaum Associates, Mahwah, NJ (August 1999).
10. *Affect in Interactions: Towards a New Generation of Interfaces*, A. Paiva, Editor, Springer Books, to appear; <http://gaiva.inesc.pt/i3ws/i3workshop.html> (1999).
11. J. D. Velasquez, "Modeling Emotions and Other Motivations in Synthetic Agents." *Proceedings of AAAI 97* (1997), pp. 10–15.
12. J. D. Velasquez, "A Computational Framework for Emotion-Based Control," Workshop on Grounding Emotions in Adaptive Systems, *Fifth International Conference on Simulation of Adaptive Behavior*, Zurich (August 1998).
13. H. Ishii and B. Ullmer, "Tangible Bits: Towards Seamless Interfaces Between People, Bits and Atoms," *Proceedings of Conference on Human Factors in Computing Systems (CHI '97)*, Atlanta, ACM Press (March 1997), pp. 234–241.
14. M. Norwood, *Affective Feedback Devices for Continuous Usability Assessment*, S.M. thesis, MIT, Media Arts and Sciences, Cambridge, MA (2000).
15. J. Scheirer, R. Fernandez, and R. W. Picard, "Expression Glasses: A Wearable Device for Facial Expression Recognition," *CHI '99 Short Papers*, Pittsburgh, PA (1999).
16. T. B. Sheridan, "Community Dialog Technology," *Proceedings of the IEEE* 63, No. 3, 463–475 (March 1975).
17. T. Starner, S. Mann, B. Rhodes, J. Levine, J. Healey, D. Kirsch, R. Picard, and A. Pentland, "Augmented Reality Through Wearable Computing," *Presence* 6, No. 4, 386–398 (1997).
18. R. W. Picard and J. Healey, "Affective Wearables," *Personal Technologies* 1, No. 4, 231–240 (1997).
19. E. Vyzas and R. W. Picard, "Online and Offline Recognition of Emotion Expression from Physiological Data," Workshop on Emotion-Based Agent Architectures at the *International Conference on Autonomous Agents*, Seattle, WA (1999).
20. J. Healey, J. Seger, and R. W. Picard, "Quantifying Driver Stress: Developing a System for Collecting and Processing Bio-Metric Signals in Natural Situations," *Proceedings of the Rocky-Mt. Bio-Engineering Symposium*, Boulder, CO (1999).
21. D. Kirsch, "The Sentic Mouse: Developing a Tool for Measuring Emotional Valence," S.M. thesis, MIT, Media Arts and Sciences (May 1997).
22. W. Ark, D. C. Dryer, and D. J. Lu, "The Emotion Mouse," *Proceedings of HCI International '99*, Munich, Germany (August 1999).
23. J. Healey and R. W. Picard, "Digital Processing of Affective Signals," *Proceedings of the International Conference on Acoustics, Speech, and Signal Processing*, Seattle, WA (May 1998).
24. J. Riseberg, J. Klein, R. Fernandez, and R. W. Picard, "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 (1998); a longer version of this paper has been accepted for publication in *Interaction with Computers*.
25. R. Fernandez and R. W. Picard, "Signal Processing for Recognition of Human Frustration," *Proceedings of IEEE ICASSP '98*, Seattle, WA (1997).
26. J. Klein, Y. Moon, and R. W. Picard, "This Computer Responds to User Frustration," *CHI 99 Short Papers*, Pittsburgh, PA (1999); a longer version has been accepted for publication in *Interaction with Computers*.
27. J. Klein, "Computer Response to User Frustration," S.M. thesis, MIT, Media Arts and Sciences, Cambridge, MA (1998).
28. J. Healey and R. W. Picard, "StartleCam: A Cybernetic Wearable Camera," *Proceedings of the International Symposium on Wearable Computing*, Pittsburgh, PA (1998).
29. J. Healey, F. Dabek, and R. W. Picard, "A New Affect-Perceiving Interface and Its Application to Personalized Music Selection," *Proceedings of 1998 Workshop on Perceptual User Interfaces*, San Francisco, CA (1998).
30. T. Marrin and R. W. Picard, "Analysis of Affective Musical Expression with the Conductor's Jacket," *Proceedings of XII Colloquium on Musical Informatics*, Gorizia, Italy (1998).
31. T. Marrin, *Inside the Conductor's Jacket: Analysis, Interpretation, and Musical Synthesis of Expressive Gesture*, Ph.D. thesis, MIT, Media Arts and Sciences, Cambridge, MA (1999).
32. K. Blocher, "Affective Social Quotient (ASQ): Teaching Emotion Recognition with Interactive Media and Wireless Expressive Toys," S.M. thesis, MIT, Cambridge, MA (May 1999).

Accepted for publication April 7, 2000.

Rosalind W. Picard MIT Media Laboratory, 20 Ames Street, Cambridge, Massachusetts 02139-4307 (electronic mail: picard@media.mit.edu). Dr. Picard is founder and director of the Affective Computing Research Group at the MIT Media Laboratory. She holds a bachelor's degree in electrical engineering from the Georgia

Institute of Technology and a master's degree and doctorate, both in electrical engineering and computer science, from MIT. She was named a National Science Foundation Graduate Fellow in 1984 and worked as a member of the technical staff at AT&T Bell Laboratories from 1984–1987, designing very large scale integration (VLSI) chips for digital signal processing and developing new methods of image compression and analysis. In 1991 she joined the MIT faculty, in 1992 was appointed to the NEC Development Chair in Computers and Communications, and in 1995 was promoted to associate professor. The author of over 70 peer-reviewed scientific articles in pattern recognition, multidimensional signal modeling, and computer vision, Dr. Picard is known internationally for pioneering research into digital libraries and content-based video retrieval. She is corecipient with Tom Minka of a best paper prize (1998) from the Pattern Recognition Society for their work on interactive machine learning with multiple models. Dr. Picard was guest editor for the special issue of the *IEEE Transactions on Pattern Analysis and Machine Intelligence* on “Digital Libraries: Representation and Retrieval,” and edited the proceedings of the First IEEE International Workshop on Content-Based Access of Image and Video Libraries, for which she served as chair. She presently serves as an associate editor of *IEEE Transactions on Pattern Analysis and Machine Intelligence*, as well as on several scientific program committees and review boards. Her recent award-winning book, *Affective Computing* (MIT Press, 1997), lays the groundwork for giving machines the skills of emotional intelligence.