Toward Machines With Emotional Intelligence

Toward Machines with Emotional Intelligence

Rosalind W. Picard
MIT Media Laboratory

Abstract

For half a century, artificial intelligence researchers have focused on giving machines linguistic and mathematical-logical reasoning abilities, modeled after the classic linguistic and mathematical-logical intelligences. This chapter describes new research that is giving machines (including software agents, robotic pets, desktop computers, and more) skills of emotional intelligence. Machines have long been able to appear as if they have emotional feelings, but now machines are also being programmed to learn when and how to display emotion in ways that enable the machine to appear empathetic or otherwise emotionally intelligent. Machines are now being given the ability to sense and recognize expressions of human emotion such as interest, distress, and pleasure, with the recognition that such communication is vital for helping machines choose more helpful and less aggravating behavior. This chapter presents several examples illustrating new and forthcoming forms of machine emotional intelligence, highlighting applications together with challenges to their development.

1. Introduction

Why would machines need emotional intelligence? Machines do not even have emotions: they don't feel happy to see us, sad when we go, or bored when we don't give them enough interesting input. IBM's Deep Blue supercomputer beat Grandmaster and World Champion Gary Kasparov at Chess without feeling any stress during the game, and without any joy at its accomplishment, indeed without feeling anything at all. Nobody worried that it would feel intimidated or get flustered. With the steady hassle of dealing with viruses, spy-ware, operating system updates, hardware failures, network delays, forgotten passwords, and a host of other computer-related ills, who wants to add worries about their machine acting in any way emotional? Just imagine that you might have to wait for it to have a certain emotional state before you could use it: "Thelma, Can you please come calm down my computer so it will let me read my email?" "Thelma, can you calm me down now too?" This scenario is so unpleasant, that a machine with a sticker labeled "emotions inside" sounds more like a joke from late-night comedy television than like anything we want to see in stores, much less in our office.

This chapter is not about giving machines emotional intelligence to make them more "emotional." Instead, it is about how emotional intelligence could address several problems that exist today, while enabling better technologies for the future. To illustrate, consider the following scenario.

Please imagine that you are the main character in the scenario, and that you are not reading this chapter, but imagine you are in your office, working very hard on a problem of great concern and importance to you. Suppose that you have an urgent deadline tomorrow, and it means a lot to you to make this deadline. Thus, you are in your office working very hard, when the following sequence of events unfolds:

> *Someone whose name you don't know enters your office. You are polite, but also slightly annoyed by the interruption, and your face is not welcoming. This individual doesn't apologize; perhaps you express a little bit of annoyance. He doesn't seem to notice. Then, he offers you advice that is useless. You express a little more annoyance. He doesn't show any hint of noticing that you are annoyed. You try to communicate that his advice is not helpful right now, but he cheerily gives you more useless advice. Perhaps when he first came in you started off with a very subtle expression of negativity, but now (you can use your imagination given your personal style) let's say, "the clarity of your emotional expression escalates." Perhaps you express yourself verbally, or with a gesture. In any case, it doesn't matter: he doesn't get it. Finally, you have to tell him explicitly to go away, perhaps directly escorting him out. Fortunately he leaves, but first he winks and does a happy little dance.*

Is this interrupting individual someone that you would eagerly invite back to your office? The annoyance he caused (when this situation happened to me) was not exactly helpful to my productivity. I do not consider his behavior to be intelligent, and he would not last long on the job if I were his boss and if this annoying behavior persisted. Nonetheless, the behavior above is wide spread – for it is that of the most successful and most famous, or infamous, computer software agent known to date, the "Clippit," the talking, dancing, smiling cartoon paperclip Office Assistant shown in Figure 1, a piece of "intelligent" software that ships standard with Microsoft Office.

Many people complain that Clippit is not very intelligent. Actually, Clippit *is* very intelligent when it comes to some things: he probably "knows" more facts about Microsoft Office than 95% of the people at MIT. He is also very good at recognizing your actions, such as whether you are writing a letter or not. He is not good at knowing what you are *intending* to do, but most people are also far from perfect at this. People don't always have the best timing when they interrupt you, and they also don't always understand your problem or how to fix it. Researchers know that recognizing intent is a hard problem for machines and they are working on it. What then is so annoying about the above scenario, and about so many people's daily experiences with this modern "intelligent" technology?

While Clippit is a genius about Microsoft Office, he is an idiot about people, especially about handling emotions. Consider three of the skills of emotional intelligence that Clippit does not have:

1. *He doesn't notice you are annoyed. [Doesn't detect or recognize your emotion]*

Even a dog can see if you are annoyed, such as when the dog is on the sofa where he knows he is not allowed, and you yell at him to get down. He notices when you yell and are upset, and he changes his behavior, putting his ears back and tail down, as well as doing the right thing by getting down (eventually). While a dog may not know your intent, he does know you are upset, and he also knows how to respond under such conditions. He responds in such a way that you want to see him again. The dog may not be as smart as Clippit when it comes to Microsoft Office, but he knows how to handle your frustration in a way that helps you feel better, not worse.

Figure 1: Microsoft Office Assitant 'Clippit.'

2. *You express more annoyance. He ignores it. [Doesn't respond appropriately to emotion]*

People don't respond to every emotion they see and neither should computers. There is a time to ignore emotion. However, there is also a time to respond, especially when the emotive individual is your valued customer, associate, friend, or family member, and when that individual is expressing increasing aggravation or annoyance toward you. To repeatedly ignore their feelings is to invite escalation of those feelings, possibly even eliciting hatred and destroying the relationship. Computers that repeatedly ignore human expressions of irritation toward them are likely to be significantly less liked as a product than those that respond intelligently to people under duress. Companies that care about their customer's feelings will pay attention to these skills in designing their technology's interactions.

3. He winks, and does a happy little dance before exiting. [Stupid about displaying emotion]

Winking, smiling, and happy dancing can be charming and entertaining in the right context. When Microsoft tested Clippit with users in its usability lab, these behaviors may have been enjoyable, especially for users who were delighted to serve as testers and felt excited about seeing a cutting-edge new technology. Clippit was probably not tested under stressful long-term conditions of personal significance to the users, taking into account variations in mood of users, i.e., under continuous real world use where feelings of genuinely significant duress occur. While it is easy to make computers look like they are having emotions (even though they are not) such as by having the machine display a happy face or dance, it is hard to give computers the ability to know when, where, and how to appropriately display emotions -- and when NOT to display them. Responding gleefully after you have annoyed somebody thrice in the last five minutes is stupid. Machines with emotional intelligence would be programmed to show emotions only if appropriate. Machines that have the potential to show emotion will also need the intelligence to know when not to show it.

The scenario above illustrates three skills of emotional intelligence with some of the problems they can cause when they are missing from technology. There are many other scenarios as well, which arise when technology interacts with medical staff, call center workers, drivers, students, and more, all in cases that impact people's feelings, productivity, and performance in measurable ways. As technology is increasingly applied to situations where it must interact with everyone – not just with its designers and creators who are experts in how it works – it is all the more vital that it does so in a way that is courteous and respectful of people's feelings.

When reading the scenario above, people tend to initially think that the intruder is a person. I deliberately worded the scenario to try to lead people to think "person" not "computer." This exercise of mentally substituting "person" for "computer" and vice-versa grew out of the work of Cliff Nass and Byron Reeves, whose book "The Media Equation" argues that human-computer interaction is inherently natural and social (Reeves and Nass 1996). While this argument may be easily believed if the computer is a humanoid robot, or perhaps a conversational software character with a face and a body, they also show that the tendency of people is to behave toward a computer as if it were a person, even if the computer has no face or life-like body, and even if it refers to its unit as "this computer" instead of "I." Their work takes studies from sociology, replaces at least one of the persons with a computer, and then tests (in a physical interaction, not just in a thought experiment as in the scenario above) if the results still hold. While the results are not always identical, the tendency to behave toward the computer *as if* it were a person is true in an enormous number of cases, confirmed not just in their work, but also by others. Thus, it can be a useful exercise when designing a new technology to ask, "What would happen if a person were to do what I'm designing this technology to do?" It is useful to pretend, for a moment, that the computer is a person. If what would happen is annoying and aggravating, then probably the technology should not be built that way.

Based on the theory of Nass and Reeves, elements of human-human interaction that are enjoyable and contribute to productivity and positive experiences have a good chance of similar outcomes when incorporated into human-computer interaction. Thus, to the extent that emotional intelligence is useful and important in human-human interaction, it is also likely to be useful and important in human-computer interaction.

2. Computers that sense and recognize emotion

Can computers know what you are feeling? I remember the look of one of my non-technical colleagues one day when she overheard some dialogue about computers that could sense people's emotion. She rolled her chair over to me, rotating so that the nearest machine was behind her back, and said to me, with a hushed but deeply worried voice, brow furrowed, and with one hand against her face to hide it from the machine and the other hand pointing over her shoulder to the machine, "Does it know that I don't like it?"

Computers do not know our innermost feelings, and there are good reasons to keep it that way, e.g., (Picard and Klein 2002), and (Reynolds and Picard 2005). Many people feel their emotions are private, and concerns about violations of privacy are every bit as valid when machines have access to our emotional information as when they have access to your social security number, banking transactions, web browsing patterns, and doctor's notes about your last visit. However, there is also a balance to be achieved, and while there are times when emotions are best kept internal, there are also times when people overtly express their feelings, and it is quite clear to everyone in the vicinity, and sometimes beyond, what they are feeling. In some cases, especially when the feelings are directed to the computer, it can be unintelligent for the computer not to recognize them. If you are cursing at the top of your lungs at the office assistant, and it doesn't flinch, but it smiles and dances and winks at you in response, then it has not only responded with a lack of emotional intelligence, it has responded stupidly. It has failed egregiously to treat its customer well. But how could it know if it can't see, hear, or otherwise sense that which is obvious to every person in the room? Computer scientists have designed technology that can know all your personal financial and health information, all your social communication, email, and entertainment preferences, your personal web browsing patterns, and more; however, they neglected to instruct the machine to see if it has upset its most important customer.

The situation could become worse when emotionally unintelligent technology is in our kitchens, cars, and living rooms, "ubiquitous," and in new forms, such as household robots. If it is smart enough to try to help us, then it should also be smart enough to adapt to our wishes, changing its behavior if we don't like it. Computers can adapt and learn how to behave in ways that we like better, but the interface for such learning and adaptation has not been a social and natural one – usually, it can only be operated by people who know a lot about computers. We are trying to change this so that machines can interpret the feedback people naturally provide to them. People might naturally praise or scold, and we are enabling the machine to see the difference. Machines can be equipped to recognize praising or scolding, or general forms of positive and negative feedback, and to respond.

There is no reason why computers have to always ignore all the feelings that are expressed at them.

What affective states should computers be enabled to recognize? While emotion theorists have developed understanding of a small set of "basic emotions" and also have developed dimensioned representations for emotion, none of the available theories provide comprehensive models for representing the states that tend to arise in interaction with computers. For example, the dominant models neglect states such as boredom, interest, and frustration. My students and I have thus not relied on these dominant theories, but have instead started with observations of a variety of interactions that occur with technology. We have been motivated by applications in learning, health, and usability, where we need to recognize what is being communicated in those contexts. Here, there is a dearth of theory, and new breakthroughs are needed in order to make useful instantiations of emotional intelligence. Our approach to developing new theory and applications is data-driven: first observe which states naturally are communicated from people to computers, then build and test models that can predict what is measured and reflected in the data. In one of our long-term projects, the building of a computerized *learning companion*, two of the key affective states found from the data are interest and boredom— two states that are not on most theorists "basic emotions" list. However, discriminating these states is vital to the machine's ability to adapt the pedagogy to help keep the learning experience going, and our model addresses them.

The Computing Research Association has listed the "building of an autonomous teacher for every learner" as one of its grand five-year research challenges. The vision for this future application is that a personalized "Socrates" or similar tutor could be available for every child to have an ideal customized learning experience. But we all know that each child is different, and a human tutor senses these differences and adapts continuously to the child. What will a learning companion need to sense? One key is sensing when to intervene in a child's learning exploration. Being able to determine the difference between students who are making mistakes while being interested and pleasurably engaged, vs. students who are making mistakes and showing increasing signs of frustration or distraction, like they are ready to quit, is really important if the system is deciding whether or not to intervene. Technology without emotional intelligence would tend to interrupt based only on the action of making a mistake, because it wouldn't be able to recognize the learner's affective state. The result is that the student who was curious and exploring may be annoyed, and thus may be discouraged from future exploration. However, if the technology waits too long before interrupting, the student who is frustrated may have already quit. What is needed is the ability to discriminate relevant emotions of the student.

While there is currently no one signal that can be read from your brain or the rest of your body to reliably tell a computer what you are feeling, there are a variety of modalities through which a computer can begin to infer information about your emotional state, especially when you are trying to communicate it clearly. The specific modalities used by a person (face, voice, gesture, etc.) may vary with context (e.g. office or home), personality (e.g. introverted or not), culture (e.g. Japanese or Italian) and other factors. We thus choose to construct a large number of tools that can be adapted to different situations.

The following are some examples of tools that we have been developing at the MIT Media Laboratory for recognizing the affective state of a person using a computer, and some of the applications that motivate their development. A discussion follows illustrating several of the current capabilities, as well as challenges researchers in this area face. The discussion is not an exhaustive one – there are many ongoing efforts and the research area is a rapidly moving one.

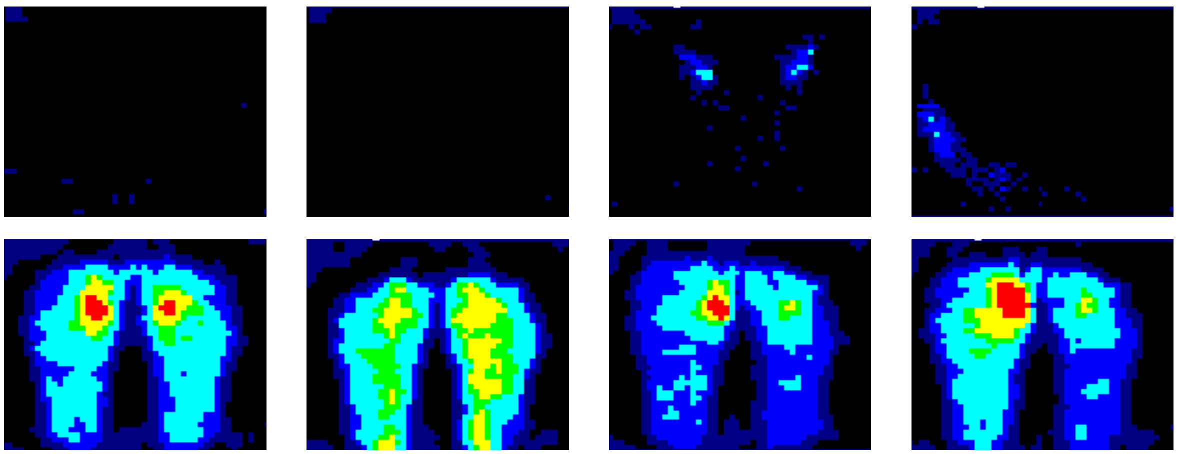

Figure 2. Pressure patterns (top row) from the back of the chair and (bottom row) seat of the chair. Left to right: sitting upright, leaning forward, leaning back, and leaning to the right.

2.1 Recognizing affective state in postural movements while seated

There's a popular belief that leaning forward is an indication of interest, while slumping lower and lower into one's chair may indicate otherwise. However, we all know that the distinction is not that simple, and there are counter-examples, such as how a student may keep leaning forward toward her computer, nudging closer and closer to the monitor, to the extreme where falls asleep with her head on the keyboard. Nonetheless, we thought that there might be some information about affective state in the posture as a function of time, so we conducted experiments to quantify any changes that might be affectively informative, developed a new tool to automate recognition of affective information from postural changes, and conducted evaluations of the tool. While the details of the work appear elsewhere (Mota and Picard 2003) the basic idea is briefly described here.

We collected data on the posture modality from children while they were seated at a computer using state-of-the-art learning software. We knew from the work of Tan et al. (Tan, Lu et al. 1997) that postures can be sensed from a commercially available set of 42 x 48 pressure sensors placed on the back of a chair and on the seat of a chair. Using custom pattern recognition software, the pressure patterns can be reliably discriminated for nine states, including the four illustrated in Figure 2: a person sitting up straight (very little action on the back of the chair, but with their weight fairly nicely distributed across the two thighs), leaning forward (no pressure on the back, more pressure toward the knees), leaning back (more color indicating pressure on the back, very little on the knees), and leaning right (more pressure on the corresponding side of the back and seat).

Selene Mota and I designed and built a new algorithm that could recognize not only nine static postures of the learner, but also analyze structure in how these postures changed as a function of time (Mota and Picard 2003). This analysis enabled the computer for the first time to use postural movements over time to classify the learner into a level of high interest, low interest, or a new state that appears to increase in frequency before boredom, which we called "taking a break." While the latter is arguably not an emotional state, we found it useful to define and recognize it as part of the challenge of recognizing the learner's affective state. While classification accuracy for the machine to recognize these three states (high interest, low interest, taking a break) was not perfect, it was significantly better than random (~33%), attaining levels from 72% to 80% accuracy, depending upon whether the system had seen the child before. With greater exposure to a person, the machine becomes more accurate (up to a point, the limits of which have not been carefully explored.)

While the chair cannot recognize your innermost feelings, it provides a modality of sensing that can be useful for the computer to know. In particular, for children seated comfortably in front of state of the art educational software, this information was found to be one of the most useful channels of feedback in discriminating their interest level (Kapoor, Picard et al. 2004), and when combined with other modalities, the accuracy of the inference improves further (Kapoor, Ahn et al. 2005).

2.2 Recognizing facial expressions

The most widely addressed area of research in automated emotion recognition, and the one where there has been the most progress, is the recognition of facial expressions. Dozens of researchers have developed methods of trying to recognize complete facial expressions such as "angry" or "sad" or to recognize individual "facial action units" that can be used to make any expression such as "inner eyebrow raise" and "eye widening" (See Pantic and Rothkrantz (2003) for a nice overview of this area of research.) Our approach has been to focus on a mixture of these, recognizing individual action units around the eyes, and coarser movements around the mouth, and to also use a special kind of camera, the IBM Blueeyes camera.

The Blueeyes camera simplifies the problem of finding the facial movements when the head is moving, by tracking the eyes. The principle used is the "red-eye effect," which takes advantage of the fact that the pupils are very good infrared mirrors, and that the pupils will thus "light up" in the image when the camera is surrounded by a ring of infrared LED's. These LED's blink on and off at the same time that two adjacent strips of LED's blink off and on, with the latter strips lighting up the face but not causing the red-eye effect. Subtracting the images where the pupils are "red" from the ones where they are not gives a quick pupil detector, which works robustly even with frequent movement in front of the camera. While the tracking fails when one blinks or when both eyes are obstructed from the view of the camera, the system recovers them very quickly as soon as they are visible again. Thus, the system is useful as the first step of a real-time facial expression recognition system.

Once the eyes are found, we apply various algorithms from computer vision and pattern recognition, which were developed in our lab by Ashish Kapoor, to try to recognize other facial movements relative to the eyes, cheeks, and mouth. While we have emphasized real-time tracking and recognition of only a subset of expressions, others have restricted head movements and allowed for some manual marking of facial features, and have attained high recognition accuracy – in the 90% range – for many facial expression components. The main problem with the methods, however, is that large and/or rapid head movements and lighting changes can ruin the ability for the system to find facial features, and thus the accuracy can drop dramatically. Our fully automatic system's accuracy, when tested on recognition of upper facial action units in the presence of a lot of natural head movements, is just under 70% accurate (Kapoor, Qi et al. 2003).

A comparison of the features of many different systems for facial expression recognition has recently been presented by Rana El-Kaliouby (el Kaliouby 2005) including many systems that claim higher accuracy rates than 70%. Most of these require some kind of manual intervention, and the rates are only reported on cases where the tracking of the facial expressions was successful. Computer vision is still far from perfect when it comes to finding and tracking faces in natural imagery where people are free to move and the sun is free to shine (or not) through the window, or outdoors, changing dramatically the appearance of the digital imagery received by the computer. Additionally, there have to date been no systems built that recognize expressions while people are moving their faces for other reasons – talking, chewing gum, bouncing to music, etc.

While research from psychology such as "Face of Interest" (Reeves 1993) suggests that faces are indicators of affective states such as interest, our findings measured from children in learning situations using the computer indicate that some children are very facially expressive while others show almost no facial expressions. Analyzing a dozen children who were using a state of the art educational program ("Fripples Place"), we found that there were no significant correlations between any of the facial features as marked by a certified facial action coding expert, and ratings of the children's levels of interest, as agreed upon by three veteran teachers. However, for some children this channel is useful, and can help the computer to be more intelligent about how it interacts with the child. Thus, one of the challenges in giving computers the ability to recognize emotion is that of knowing when a channel's lack of signal is meaningful or not.

Stronger results can sometimes be obtained by combining multiple channels, e.g. face with voice (Huang, Chen et al. 1998), face with chair, or mouse and task (Kapoor, Ahn et al. 2005). Different kinds of sensing (use of a chair, mouse, camera, etc.) may be more or less natural in different kinds of environments, and confidence that the computer has properly recognized the person's affective state tends to increase with more than one mode of sensing. There is also new pioneering research on automated recognition of facial expression and on head gestures to discern states such as "concentrating," "disagreement," "thinking," "unsure," and "interested,"(el Kaliouby and Robinson 2005; el Kaliouby and Robinson 2005), all of which could also be helpful to a machine trying to appear more caring by adjusting its responses to those of the person with whom it is interacting. El-Kaliouby has additionally described potential applications of her algorithms to assist people with autism spectrum disorder, who routinely have difficulty recognizing others cognitive-emotional states (el Kaliouby and Robinson 2005).

2.3 Recognizing emotion from physiology

In some contexts, e.g. medical monitoring or use of other wearable systems, it can be natural and comfortable to work with physiological sensors, sensing information such as changes in heart rate, muscle tension, temperature, skin conductance, and more. For example, if a patient is wearing a heart monitor, for health monitoring or for fitness tracking, then the technology can potentially measure heart rate variability changes associated with cognitive and emotional stress as well. Physiological information has been shown to carry information that changes with different emotions (Ekman, Levenson et al. 1983) and such information can be used to build classifiers for an individual's affective state. One of the challenges in building a physiology-based emotion recognition system is that emotion is only one of the factors that affects physiology: Diet, exercise, sleep, basic personality characteristics, environmental context, activity, and more also affect physiological changes, and separating out the effects due to emotion involves significant technical hurdles.

We designed, built, and tested what appears to have been the first physiology-based emotion recognition system for recognizing an individual's emotions over time. This system used four physiological signals that learned patterns for an individual over many weeks, and then achieved 81% recognition accuracy classifying which one of eight states (anger, joy, sadness, hatred, platonic love, romantic love, reverence, and neutral) that the individual was having (Picard, Vyzas et al. 2001). (The person was seated in a relatively quiet space, and deliberately focusing on having each of the eight emotions.) While the results of physiology-based recognition are promising, this work remains very technically challenging and there is a lot of physiological understanding that still needs to be developed before large repertoires of emotion can be sensed. Also, there will probably be limits regarding how far this work will scale – at some level the physiological differentiation in emotions may become minute, and possibly only available at a bio-chemical level that also involves knowing spatial locations of such activity. Sensors of such information are far from portable and can be very invasive. This is a wide-open area of research, which will need to be informed by cultural, social, and ethical guidelines as well as by scientific inquiry. There are probably certain levels of feelings that it is best to keep from revealing to others, including your personal computer.

Signals such as skin conductance and heart rate variability have also been shown to be indicators of stress in natural situations, e.g., measured from people while they were driving in Boston (Healey and Picard 2005). Recently it has also been shown that a computer software agent's empathetic responses can influence skin conductance in a way that is associated with decreased stress (Prendinger, Mori et al. 2005). In the latter case, physiological sensing of affect provides evidence that some computer behaviors can be less stressful than others. In the future, computer applications may choose behaviors based on your personal response – choosing to adapt in ways that reduce your stress, instead of always doing the same thing for everyone, regardless of the stress it causes them.

2.4 What is a natural set of sensors?

Lots of factors contribute to deciding whether a sensor of emotion is perceived as natural, comfortable, intrusive, or otherwise acceptable or not. While sometimes using an explicit sensor (such as a physiological sensing system, with skin-surface contact sensors) is valuable, say for knowing that you are being sensed, and for giving you control over when this does and does not happen, it is also the case that such an apparatus can require extra effort, and thus can be seen as "too much work to use." Some people see any kind of visible sensor as intrusive. In a medical context, however, where such sensors might naturally be present for other useful functions, they are less likely to be seen as intrusive. Sometimes it is desirable to integrate sensors into the natural environment and see what can be learned about emotions from information they gather. However, sometimes such sensors can be seen as even more invasive, since they can gather information without your knowledge.

A compromise is to make sensors that are visible so that you can tell they are there, but also to integrate them naturally into the environment, so that the person using them doesn't have to do anything special (like attaching electrodes) to communicate through the sensors. The chair sensor described earlier, falls into this category. Similarly, IBM researchers placed physiological sensors into computer mice (Ark, Dryer et al. 1999) and used a function of those sensors to classify several basic emotions. While the mouse is comfortable in many ways, our own experience with the IBM mouse was that it picked up a lot of motion artifacts on the electrodes placed on its surface, but that the varying pressure applied to the mouse might itself be useful as a sign of stress or frustration for many users.

Carson Reynolds designed, built, and tested a pressure sensitive mouse, examining differences in pressure among users doing the same task, with low stress and with slightly higher stress, and finding some preliminary evidence of stress-related pressure changes (Reynolds 1999). Inspired by this work, Jack Dennerlein at the Harvard School of Public Health decided to look more carefully at stress induced by mild occurrences of bad usability while people filled out forms on the computer. We helped Dennerlein automate his health demographics intake form, so that people had to use the mouse to fill it out. Also, unbeknownst to the subjects, we made the forms slightly annoying to fill out, with subtle bad usability in the form's design, such as the use of pull-down menus, and requiring commas between numbers. Some people found the forms to be very frustrating to fill out, while others did not find them frustrating at all. Comparing the most frustrated top half to the bottom half (according to people's self-reported frustration level), the top half was found to apply significantly more force to the mouse than the bottom half, especially right after being notified that they made an error on the previous page, and had to go back and fix it. When they went back to fix it, they found that it had erased all of their answers on that page and they had to reenter them. Analyzing the data later, Dennerlein found that frustration did co-occur with greater force, and also the specific muscles activated were those causally related to wrist injuries (Dennerlein, Becker et al. 2003). This finding suggested a new application: a mouse that "squeaks" or notifies you in a subtle way when you might be pressing it harder than necessary for some length of time. Thus, the technology could help you become aware of behaviors, possibly linked to stress, which are also associated with medical injury, so that you could better regulate those behaviors and their associated feelings before such an injury develops.

The mouse is just one example of placing affective sensing into devices that we are in ordinary physical contact with. Steering wheels, telephones, keyboards, and more, present opportunities for computers to make more accurate inferences not only about what a person is trying to do, but also *how* smoothly (or not) the task is progressing. While none of these techniques allows a computer to know your innermost feelings, they are channels that can be used to help you communicate a few of the adverbs that might help a machine better customize its interaction with you.

2.5 Sensing emotion from dialogue

At times it is appropriate to just ask somebody, "How are you feeling?" While somebody may say, "I'm fine" (even when they are clearly miserable) the self-report of feelings is the de facto standard among psychology researchers when it comes to assessing emotion. Psychologists routinely give people forms to fill out, where they check off items on seven-point scales, or use other expressive metrics such as the Self-Assessment Manikin (Bradley and Lang 1994). Machines can also ask people what they are feeling on a questionnaire, or they can do so in a more conversational way, as indicated in the simple dialogue boxes shown in Figure 3.

In the dialogue in Figure 3, we aim to go beyond just collecting self-report data. Here, the computer also responds to the user's statement in a way that tries to show the machine is "listening" and "trying to understand" what the user has said. The machine does this by attempting to paraphrase the user's response, and by giving the user a chance to correct it if the paraphrase is wrong: to say "actually the feelings were a little worse than that," or "a little better." Thus, the machine not only collects the user's self-reported feelings, but it also attempts to help the user manage negative emotions by empathizing when things are not going well.

The technique in Figure 3 was part of the first "empathetic" agent built by Jonathan Klein. After measuring the behavior and responses of 66 people who were either users of this agent or users of one of two controls, the agent's empathetic approach was argued to have a positive influence on user's behavior. The controls included a similar dialogue-based agent that simply ignored people's emotions, and another one that let people vent about their emotions but did not otherwise respond to them (Klein, Moon et al. 2002).

While machines remain very limited in their ability to understand most of language, they can already engage successfully in quasi-scripted dialogues about feelings (Klein, Moon et al. 2002; Bickmore and Picard 2005; Liu and Picard 2005). The dialogues can reveal feelings in a way that is natural and comfortable for people, e.g.,

> Computer: "How's it going?"
> Person: "Not so great."

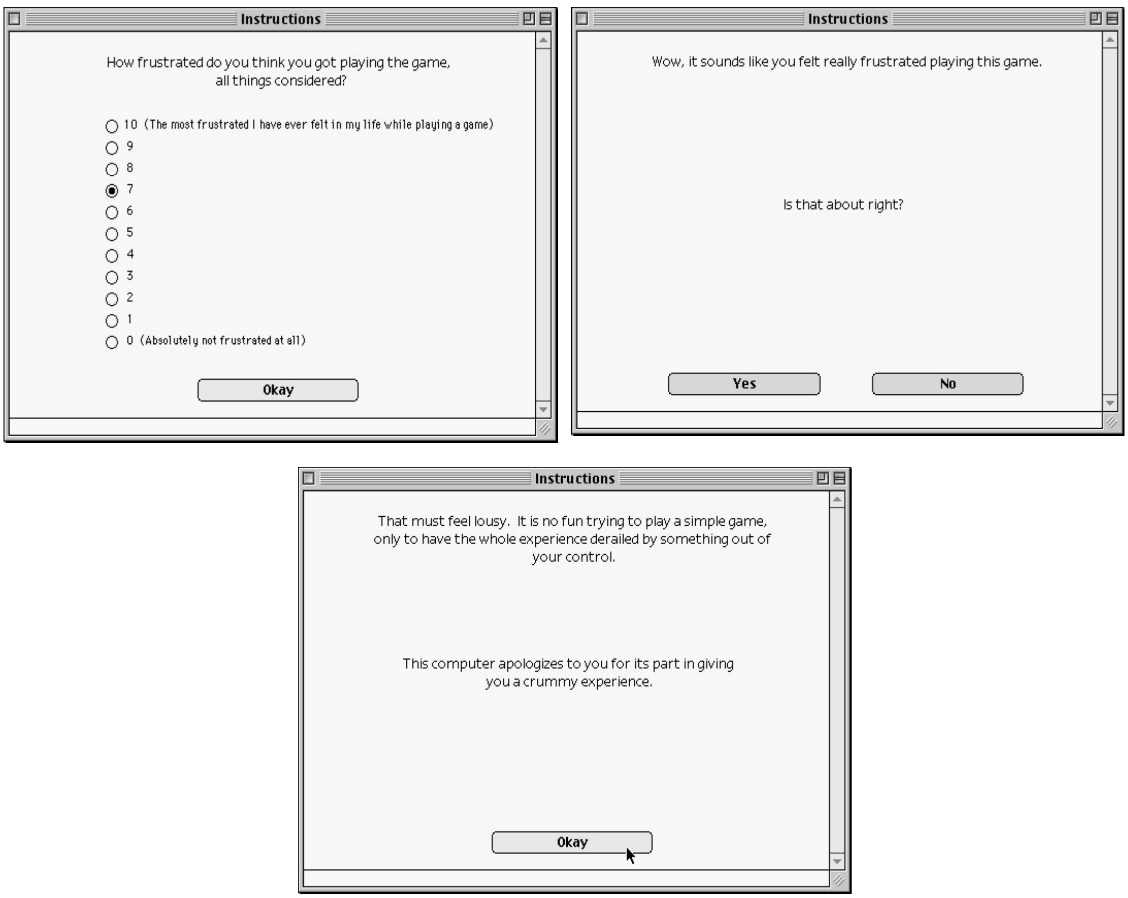

Figure 3: Klein et al. used simple dialogue boxes to ask about emotion, and to convey the impression of active listening, empathy, and sympathy to help frustrated computer users feel better.

> Computer: "Oh dear. Sorry to hear. Anything I can do to help?"
> Person: "I don't know, I just feel terrible about …"

Dialogue systems have a chance to sense words selected by a person and reason about the associated affect using either classic rule-based reasoning tools from Artificial Intelligence or using new "common sense reasoning" tools (Elliott 1992; Liu, Lieberman et al. 2003; Ma, Osherenko et al. 2005). Using new tools of the latter ilk, Shaundra Daily has recently built a new technology that tries to help teen girls reason about the emotion of themselves and others in their life (Daily 2005). The tool helps girls compose and illustrate stories about events of personal significance, and the common sense reasoning makes suggestions about possible emotions that may have been present in the events the girl is writing about. After the system suggests "Perhaps this made you feel angry?" if the girl says, "yes, that's how I felt", then the system also has the opportunity to subtly empathize, helping validate the girl's feelings as well as reinforce her reflection on them. An initial study of use of this system by girls in a poor-performing school showed that their incorporation of emotional words in their stories increased over time, compared to a control group who used the story telling system without the computer's emotion suggestions (Daily 2005). This particular application of the technology also illustrates how even when the technology cannot always accurately infer the true feelings, it can still offer suggestions in a considerate way, which may in turn lead to productive outcomes.

Sometimes feelings are not communicated through *what* is said as much as *how* it is said, e.g. "Good Morning!" can be spoken out loud with genuinely cheery enthusiasm or with annoyance, disdain, and other kinds of inflection that may clearly belie the meaning of the words. Machines also have the opportunity to listen to para-linguistic aspects of speech for indications of a person's feelings. Emotion recognition in speech is an active area of research (Douglas-Cowie, Cowie et al. 2003; Fernandez and Picard 2005).

Systems that can intelligently sense the affective intent of language can be of immediate use in call centers and in other applications where the intelligent management of emotion has important business implications, as well as implications for managing the health of the workers who take the calls. Call centers tend to have high turnover of staff, and their cost to companies is enormous in continuously training and managing these stressful enterprises. If the worker were to have an empathetic system warning her, "This one sounds like it may be tough, you might want to take a deep breath first…" and showing concern for her feelings afterward if it looks like it may have indeed been stressful, "Was that caller as difficult as I guessed?" and if she answers "Yes, almost the worst I've ever had" then it's time for a little empathy, and perhaps even some congratulations to her or some encouragement for maintaining her own composure. While managing feelings is no substitute for skilled training in how to handle a customer's mishap, skilled training alone is also not enough. Fixing the problem without regard for how it has made the customer feel can leave a customer still dissatisfied, while sometimes even when a problem is beyond fixing, a customer can feel like the company cares and did their best, engendering loyalty. As technology becomes a greater part of the customer-service interaction, it too must be taught about this important dance between fixing problems and managing feelings.

3. Computers that respond appropriately to human emotion

Computers usually ignore emotion. Most of them can't sense it, and even if they can sense some aspects of it, the rates are still not perfect, and the theory of emotion is still such a poor fit to reality, that it can be risky to assume too much in responding. However, in some cases there are appropriate ways that computers can respond, and there is evidence that it can be productive to do so.

In our work we have been particularly concerned with how to respond to emotions that occur most frequently around computers. In current business and home uses, these include emotions such as frustration, distress, and even anger. A survey by Mori in the United Kingdom, entitled "Rage Against the Machine" found that of 1250 workers questioned in the UK, four out of five have "seen colleagues hurling abuse at their personal computers," three quarters admit that they swear at their computers, and nearly half of all people working with computers feel frustrated or stressed because of information technology problems. Lest people think that the current generation is overreacting, and these problems will disappear with the next generation of youth, they also found that a quarter of all users under 25 admit they have kicked their computer (1999).

Suppose that a computer infers with high probability that you are frustrated. When should it ignore this and when should it respond? If it should respond, then how should it respond? Handling of emotions, in others as well as in oneself, involves emotional intelligence. This is a difficult challenge for many people, and even more so for machines. Machines have a very long way to go before they can recognize emotions as well as people, much less be smart about how they can respond to them without causing negative emotions to escalate. Nonetheless, the topic is an important one to tackle, and successful handling of emotion is critical to good customer service.

We have begun to give computers a limited range of abilities when it comes to helping handle the emotions of people with whom they are interacting. Our first efforts in this area have been to respond to negative emotions with a little bit of active listening, empathy, and sympathy. As illustrated in Figure 3, these responses do not have to involve faces, voices, or other visual effects (dancing, etc.) It is also the case that the computer need not refer to itself as having feelings or even refer to itself as "I" (note that there is no embodied character and no "self" referencing in the empathetic dialogues by Klein et al.)

Brave, Nass, and Hutchinson compared the use of empathic facial displays and text messages by an embodied computer agent (using images of a person's face with different emotional displays) with self-oriented emotional displays and messages (Brave, Nass et al. 2005). They found that the empathic agent was given more positive ratings, including likeability and trustworthiness, as well as greater perceived caring and felt support, compared to either an agent that used self-oriented displays and messages or an agent that performed no emotion-oriented behavior. Drawing from human-human interaction, such findings are no surprise. Comparing the potential applications in human-computer interaction to existing ones in human-human interaction, we find that in many cases there are important outcomes related to use of empathy. For example, in physician-patient interactions, the physician's empathy for a patient plays a significant role in prescription compliance, and a physician's *lack* of empathy is the most frequent source of complaints (Frankel 1995). A variety of computer scientists and interaction designers have started to imbue computers with various empathetic and even "caring" behaviors, especially for applications in health and education (Lisetti, Nasoz et al. 2003; Paiva, Dias et al. 2004; Liu and Picard 2005; Prendinger and Ishizuka 2005; Prendinger, Mori et al. 2005).

3.1 Relational Computers

Affective responses play an important role in long-term relationships between people, and have recently become a topic of study in the development of "Relational Agents", computer characters that are programmed to develop long-term social-emotional relationships with people (Bickmore 2003). These agents would have application in areas such as health maintenance or addiction treatment, where long-term compliance is improved by ongoing interaction with an expert helper (and where spiraling health care costs make it out of the reach of most people to have such a full-time helper).

In one study, an embodied conversational agent, "Laura," was developed to assist people who wanted to be walking thirty minutes a day, and weren't succeeding yet at this goal, yet knew it was important to get exercise. Two versions of the agent were built: one that was friendly, social in its greetings (and farewells), and talked with you about your goals and progress, trying to motivate you in ways that were derived from observations of a human fitness trainer, and a second (the condition of most interest to us) that was identical to the first, but that also tried to develop a social-emotional long-term relationship with you. The latter was always happy to see you, and after some initial interaction it would begin its sessions by asking how you were feeling and responding empathetically with both words and appropriate facial expressions. It also would "move closer" (the animation was zoomed) when you talked about personal topics such as about your feelings, and over time it changed its language in subtle ways, similar to how people change theirs as they get to know you. (Details of all the differences can be found in Bickmore's Ph.D. thesis (Bickmore 2003)). Subjects were randomly placed into a group that interacted with the non-relational or the relational Laura approximately daily on their home computers for one month, during which time Laura "chatted" with them for about five minutes, giving feedback on their exercise behavior, discussing any obstacles they mentioned about exercise, providing educational content related to exercise, and obtaining and following up on commitments to exercise. The task and goal structure of the dialogues were the same across the two conditions, as was the interface and the animations of Laura.

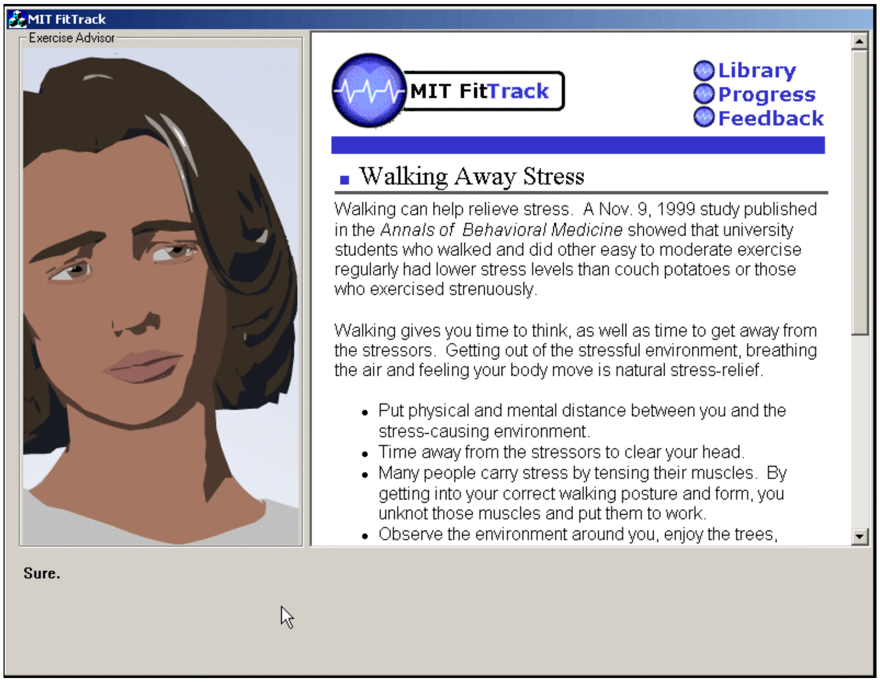

Figure 4. Relational agent, "Laura" displaying a look of concern after a user expresses feeling badly.

About a hundred people were recruited for a six-week study in which 1/3 interacted with non-relational Laura, 1/3 interacted with relational Laura, and 1/3 interacted with a no-agent control. In all cases, subjects were encouraged to login to the system daily (the average rate of login was every-other-day) and report on their activity, pedometer readings, and (in the two agent conditions) have a short chat with the agent. The principal outcome measure in comparing the relational and non-relational agent conditions was the Working Alliance Inventory, a 36-item self-report questionnaire used in psychotherapy that measures the trust and belief that the therapist and patient have in each other as team-members in achieving a desired outcome (Horvath and Greenberg 1989). This inventory has three subscales: Bond, Task, and Goal. While there were no statistically significant differences in the Task and Goal measures after one week and after one month of interaction with Laura, there were important and statistically significant differences in the Bond subscale. This measure is of particular interest in thinking about emotional intelligence, as it assesses the emotional bond between the helper (software agent) and the helpee (person).

Thirty-three subjects completed the month of interactions with the relational agent and twenty-seven subjects completed interactions with the non-relational agent. Subjects were mostly (69%) students and were 60% female (balanced across the two conditions). While a full description of the results (including exercise-related outcomes) can be found elsewhere (Bickmore 2003), key findings relating to emotional intelligence included significant differences in perception that Laura liked them $t(60)=2.56$, $p< 0.05$, in trust of Laura and feeling trusted by Laura $t(50)=2.05$, $p< 0.05$, in feeling Laura was concerned about them $t(60)=2.19$, $p<0.05$, and in feeling Laura cared about them ("in her own unique way", a clause which was added so that people would not feel entirely silly answering these questions about a software agent that had told them on day one that it was just a computer program) $t(60)=2.39$, $p<0.05$. When asked at the end of the month if they would like to continue working with Laura, subjects in the relational condition also responded much more favorably than the non-relational group, $t(57)=2.43$, $p=.009$. This measure is of particular importance since the aim is long-term interaction with a technology, and in health related applications, continuing with a treatment program is related to outcome, and desire to continue with a treatment is likely to facilitate that result as well.

In this study, we also obtained a behavioral measure while assessing subjects' feelings of caring toward the agent. Throughout the daily interactions, the agent spoke to the subject, while the subject clicked on text responses to talk to the agent. In the closing session, the text responses for the subjects to bid farewell comprised two choices: "Bye" and "Take care Laura, I'll miss you." Significantly more subjects in the relational group (69%) chose the most sentimental farewell ("Take care Laura, I'll miss you.") than in the non-relational condition (35%), $t(54)=2.80$, $p=.004$. This behavioral finding is consistent with the interpretation that the relational agent was more successful in eliciting a sense of bond and caring in the people interacting with it. Thus, it suggests that the skills given to the relational agent may bode well for maintaining longer term interactions than when the agent was lacking relational skills (the non-relational condition).

5. Summary and the Future

Machines are beginning to receive abilities to sense and recognize human emotion, to respond to it with better skills so that negative feelings are less likely to escalate, and to have other skills that can be used to help people develop and assess various abilities and behaviors that contribute to emotional intelligence. These skills are changing the way applications interact with people – agents will show consideration of human feelings and take steps to be less annoying, and applications that deal with customers will show respect for their feelings by acknowledging them and adjusting their behavior in response. The rates at which people curse at their machines, report observations of "colleagues hurling abuse at their machines," and kick their machines, should all drop. Experiences with information technologies should become less frustrating and more productive.

A variety of new applications will be enabled –technology that helps aid individuals in reflection on emotions, and in learning about and developing emotional skills, as well as technology that helps assess such skills.

Motivated by progress building an Affective Learning Companion, we have recently begun considering a project to build an Assessment Companion for Emotional Intelligence. This system would target the assessing of how people *behave* in situations that demand emotional intelligence, vs. testing what they cognitively infer about such situations as in current tests of emotional intelligence such as the MSCEIT (Mayer, Salovey et al. 2003). This distinction between knowing and doing, with respect to emotion in cognition, was illustrated by Damasio's patients (Damasio 1994), who had frontal lobe and other damage to emotional areas in their brain, and yet were able to reason fine about emotional situations. While they could score fine on pencil-and-paper tests, the situation was quite different when it came to how they behaved. When these patients had to act appropriately, or do the right thing, in real-time, in complex social situations, their behavior was far from normal. While their cognitive reasoning about emotion was fine, their ability to use emotion in action selection was impaired.

Scientists could begin to build interactive experiences (e.g. virtual or virtually augmented environments, social games, and more) that could put emotional intelligence to tests involving behavior, not merely cognitive reasoning. These environments would be engaging, unlike the taking of a long test, which we found to be a problem with our teen subjects in one experiment (Daily 2005), where they indicated that the test was much too long and boring for them to take it seriously.

While some future testing environments could be built without giving machines emotional abilities, rich assessments with computer role-playing and more could be constructed if the technology could also sense and respond to certain kinds of emotional cues. Thus, giving machines emotional sensing could help them to assess human nonverbal behaviors, and respond to these in ways that "push the boundaries" of people's feelings, allowing for a rich exploration of how they behave in different emotional states. Such machines might have goals quite opposite to those that we are giving to office equipment: In the latter we aim to reduce frustration, but the assessing companion could aim to maximize it, and test how an individual manages himself or herself in the face of difficulty. The computer could deliberately provoke frustration and irritation, and examine its influence on the person's ability to perform under stress.

Our machines will undoubtedly continue to improve in their ability to calm, comfort and soothe us, to earn our trust and rapport, and to help us in a greater variety of situations. Machines can make use of increasingly sophisticated techniques in how they address our problems. For example, rather than just offering canned empathetic messages or messages indexed to degree of emotional upset as in (Klein, Moon et al. 2002; Bickmore and Picard 2005; Brave, Nass et al. 2005; Bickmore, Gruber et al. to appear), they can offer comforting messages that are formulated according to a more fine-grained set of emotional criteria, such as: extent to which a user's feelings are explicitly acknowledged, elaborated and legitimized; extent to which the messages are centered on the user's emotions (vs. the computer's feelings or the causes of the upset); and whether the empathic messages also contain a cognitively-oriented explanation of the user's emotions or not (Burleson 1985; Burleson and Picard 2004). These interventions can also help users re-examine the events that gave rise to their negative emotions in the first place.

The aim is not just to build machines that have emotional intelligence, but to build tools that help people boost their own abilities at managing emotions, both in themselves and in others.

Acknowledgments

I would like to thank Shaundra Bryant Daily for helpful comments on this manuscript. This research was supported in part by the MIT Media Lab's Things That Think Consortium, Digital Life Consortium, and by the National Science Foundation ITR 0325428. Any opinions, findings and conclusions or recommendations are those of the author and do not necessarily reflect the views of the NSF.

References

(1999). Business: Costly computer rage. BBC News.

Ark, W., D. C. Dryer, et al. (1999). The Emotion Mouse. HCI International '99, Munich, Germany.

Bickmore, T. (2003). Relational Agents: Effecting Change Through Human-Computer Relationships. Media Arts and Sciences. Cambridge, MIT.

Bickmore, T., A. Gruber, et al. (to appear). "Establishing the computer-patient working alliance in automated health behavior change interventions." Patient Educ Couns.

Bickmore, T. and R. Picard (2005). "Establishing and Maintaining Long-Term Human-Computer Relationships." ACM Transactions on Computer Human Interaction **12**(2): 293-327.

Bickmore, T. and R. W. Picard (2005). Future of Caring Machines. Future of Intelligent and Extelligent Health Environment. R. Bushko, IOS Press. **118 Studies in Health Technology and Informatics**.

Bradley, M. M. and P. J. Lang (1994). "Measuring emotion: the Self-Assessment Manikin and the Semantic Differential." Behav Ther Exp Psychiatry **25**(1): 49-59.

Brave, S., C. Nass, et al. (2005). "Computers that care: investigating the effects of orientation of emotion exhibited by an embodied computer agent." Int J Human-Computer Studies **62**: 161-178.

Burleson, B. (1985). "The Production of Comforting Messages: Social-Cognitive Foundations." J Lang and Social Psyc **4**(3&4): 253-273.

Burleson, W. and R. W. Picard (2004). Affective Agents: Sustaining Motivation to Learn through Failure and a State of Stuck. Social and Emotional Intelligence in Learning Environments: Workshop In conjunction with the 7th International Conference on Intelligent Tutoring Systems, Maceio-Alagoas, Brazil.

Daily, S. B. (2005). Digital Story Explication as it Relates to Emotional Needs and Learning. Media Arts and Sciences. Cambridge, MIT.

Damasio, A. R. (1994). Descartes' Error: Emotion, Reason, and the Human Brain. New York, Avon Books.

Dennerlein, J., T. Becker, et al. (2003). Frustrating Computer Users Increases Exposure to Physical Factors. International Ergonomics Association, Seoul, Korea.

Douglas-Cowie, E., R. Cowie, et al. (2003). Special Issue on Speech and Emotion. Speech Communication. **40**.

Ekman, P., R. W. Levenson, et al. (1983). "Autonomic Nervous System Activity Distinguishes Among Emotions." Science **221**: 1208-1210.
Elliott, C. (1992). The Affective Reasoner: A process model of emotions in a multi-agent system, Northwestern University.
Fernandez, R. and R. Picard (2005). Classical and Novel Discriminant Features for Affect Recognition from Speech. Interspeech 2005 - Eurospeech - 9th European Conf on Speech Communication and Technology, Lisboa, Portugal.
Frankel, R. (1995). "Emotion and the Physician-Patient Relationship." Motivation and Emotion **19**(3): 163-173.
Healey, J. and R. W. Picard (2005). "Detecting Stress During Real-World Driving Tasks Using Physiological Sensors." IEEE Trans. on Intelligent Transportation Systems **6**: 156-166.
Horvath, A. and L. Greenberg (1989). "Development and Validation of the Working Alliance Inventory." Journal of Counseling Psychology **36**(2): 223-233.
Huang, T. S., L. S. Chen, et al. (1998). Bimodal emotion recognition by man and machine. ATR Workshop on Virtual Communication Environments.
el Kaliouby, R. (2005). Mind-Reading Machines: Automated Inference of Complex Mental States. Computer Laboratory, University of Cambridge.
el Kaliouby, R. and P. Robinson (2005). "The Emotional Hearing Aid: An Assistive Tool for Children with Asperger Syndrome." Universal Access in the Information Society **4**(2).
el Kaliouby, R. and P. Robinson (2005). Real-time Inference of Complex Mental States from Facial Expressions and Head Gestures. Real-Time Vision for Human-Computer Interaction, Springer-Verlag**:** 181-200.
Kapoor, A., H. Ahn, et al. (2005). Mixture of Gaussian Processes for Combining Multiple Modalities. Proceedings of the Multiple Classifier Systems, 6th International Workshop, MCS 2005, Seaside, CA.
Kapoor, A., R. W. Picard, et al. (2004). Probabilistic Combination of Multiple Modalities to Detect Interest. International Conference on Pattern Recognition, Cambridge, U.K.
Kapoor, A., Y. Qi, et al. (2003). Fully Automatic Upper Facial Action Recognition. IEEE International Workshop on Analysis and Modeling of Faces and Gestures (AMFG 2003) held in conjunction with ICCV 2003, Nice, France.
Klein, J., Y. Moon, et al. (2002). "This Computer Responds to User Frustration: Theory, Design, Results, and Implications." Interacting with Computers **14**: 119-140.
Lisetti, C., F. Nasoz, et al. (2003). "Developing multimodal intelligent affective interfaces for tele-home health care." Int J Human-Computer Studies **59**(1-2): 245-255.
Liu, H., H. Lieberman, et al. (2003). A Model of Textual Affect Sensing using Real-World Knowledge. International Conference on Intelligent User Interfaces, Miami, Florida.
Liu, K. K. and R. W. Picard (2005). Embedded Empathy in Continuous Interactive Health Assessment. CHI Workshop on HCI Challenges in Health Assessment, Portland, OR.
Ma, C., A. Osherenko, et al. (2005). A Chat System Based on Emotion Estimation from Text and Embodied Conversational Messengers (Preliminary Report). 2005 IEEE Int'l Conf on Active Media Technology (AMT-05), Takamatsu, Kagawa, Japan.

Mayer, J. D., P. Salovey, et al. (2003). "Measuring Emotional Intelligence With the MSCEIT V2.0." Emotion **3**: 97-105.
Mota, S. and R. Picard (2003). Automated Posture Analysis for Detecting Learner's Interest Level. Workshop on Computer Vision and Pattern Recognition for Human-Computer Interaction, CVPR HCI, Madison, WI, IEEE.
Paiva, A., J. Dias, et al. (2004). Caring for Agents and Agents that Care: Building Empathic Relations with Synthetic Agents. Proceedings of the Third International Joint Conference on Autonomous Agents and Multi Agent Systems (AAMAS-04), New York, NY, ACM Press.
Pantic, M. and L. J. M. Rothkrantz (2003). "Toward an affect-sensitive multimodal human-computer interaction." Proc. of the IEEE **91**(9): 1370-1390.
Picard, R. and J. Klein (2002). "Computers that recognize and respond to user emotion: theoretical and practical implications." Interacting with Computers **14**: 141-169.
Picard, R. W., E. Vyzas, et al. (2001). "Toward Machine Emotional Intelligence: Analysis of Affective Physiological State." IEEE Transactions Pattern Analysis and Machine Intelligence **23**(10).
Prendinger, H. and M. Ishizuka (2005). "The Empathic Companion: A Character-based Interface that Addresses User's Affective States." Int'l J of Applied Artificial Intelligence **19**(3-4): 267-285.
Prendinger, H., J. Mori, et al. (2005). "Using Human Physiology to Evaluate Subtle Expressivity of a Virtual Quizmaster in a Mathematical Game." Int'l J of Human-Computer Studies **62**: 231-245.
Reeves, B. and C. Nass (1996). The Media Equation. New York, Cambridge University Press.
Reeves, J. M. (1993). "The Face of Interest." Motivation and Emotion **17**(4): 353-375.
Reynolds, C. (1999). Measurement of Frustration with Computers. Media Arts and Sciences. Cambridge, MIT.
Reynolds, C. and R. W. Picard (2005). Evaluation of Affective Computing Systems from a Dimensional Metaethical Position. 1st Augmented Cognition Conference, In conjunction with the 11th International Conference on Human-Computer Interaction, Las Vegas, NV.
Tan, H. Z., I. Lu, et al. (1997). The Chair as a Novel Haptic User Interface. Proceedings of the Workshop on Perceptual User Interfaces, Banff, Alberta, Canada.