Ethical Evaluation of Displays that Adapt to Affect

Carson Reynolds
MIT Media Laboratory
20 Ames St, Room E15- 120F
Cambridge, Massachusetts 02139
carsonr@media.mit.edu

Rosalind Picard
MIT Media Laboratory
20 Ames St, Room E15- 00G
Cambridge, Massachusetts 02139
picard@media.mit.edu

Abstract:
We are developing new kinds of sensors and algorithms that can sense and begin to interpret information related to affect. Starting from the standpoint of Value Sensitive Design, which considers values such as calmness, autonomy, and informed consent, displays that adapt to affect can be ethically evaluated. The question of how adaptive technologies act on these values is of critical importance to the affective computing group. We are seeking to better understand ethical implications for designers through empirical experiments.

Suppose you are a soldier faced with the following dilemma. You've been issued a prototype HUD while acting as team leader of an infantry rifle squad. The HUD prioritizes communications based on the sender's emotional state. You are told this will allow you to focus on squad members in distress and in need of the most assistance. However, when you wear the HUD, the constant fluctuation of message priority (which mirrors constant changes in emotion) makes you disoriented. Furthermore, your squad is performing overwatch, and you are expected to remain extremely vigilant for contact with the enemy. So the dilemma facing you is: should you wear the HUD or not? Given the HUD's impact on your mission but potential improvement of your ability to communicate with your squad, this may be a difficult decision to make.

Adaptive displays face problems of predictability and controllability [Jameson, 2002], especially when fed with emotional inference data that may itself vary unpredictably over time. Adaptive algorithms seek out what is suitable, given some notion of what is optimal [Reynolds, 2002]. With adaptive displays what is optimal often has to deal with emotional, cognitive, or physical workload.

But the dilemma facing the soldier in the example above can also be thought of in terms of ethical suitability. Considering the good and harm of different choices in a dilemma may itself be an interesting adaptive strategy.

Some might say that the soldier's job isn't to consider ethics. The soldier's job is instead to attrite the enemy [Peters, 2004]. Furthermore, officers already make use of a Code of Conduct [Department of Defense, 1998].

Still, in the process of following orders a soldier confronts dilemmas [Department of the Army, 2003]. Researchers of adaptive display technologies that could be used by soldiers may find it informative to consider these dilemmas as well.

In the particular case of adaptive displays dealing with emotion the Army's Soldier's Guide describes one step in the evaluation of different courses of action (COA) in the following manner:

> "Gut check"—Does the COA "feel" like it is the right thing to do? Does it uphold Army values and develop your character or virtue? [Department of the Army, 2003]

The Affective Computing group specializes in sensors and algorithms whose goal is to recognize and communicate emotional responses [Picard and Klein, 2002]. The application of this technology to adaptive displays for soldiers would yield devices that could track the affective responses of their squad. For instance a HUD display that is fed by sensors and algorithms that look for signs of interest, alertness, stress, and possible fatigue (among other states).

Today, these sensors and algorithms operate with less- than perfect accuracy. The estimates of emotional states are usually probabilistic and less accurate than the observations of a skilled person, although they tend to be much better than random guessing. The algorithms for estimating state require significant processing and in some cases can provide results in real time [Kapoor and Picard, 2002].

The Affective Computing group has been approaching adaptive systems with two prongs: algorithms and sensors. Algorithmically speaking, we have developed new machine learning algorithms that are suitable in application to recognition of emotional states [e.g. Qi and Picard, 2002]. These algorithms differ from more traditional adaptive algorithms like Genetic Programming [Goldberg, 1989], Simulated Annealing [Kirkpatrick et al., 1983], or Conjugate Gradient Decent [Shewchuk, 1994] because they focus on recognition instead of optimization. These recognition algorithms are also fed with a variety of sensors that are tailor- made to suit the collection of signals relating to emotion states like frustration. For instance a pressure- sensitive mouse has been designed to record hand force on the mouse surface [Reynolds, 2001].

These sensors and algorithms can be fused to provide systems that act as affective classifiers, whose goal is to provide reliable estimates of emotional state. Hooking a classifier into an adaptive display then becomes relatively straightforward. The labels from the classifier are next fed into the suitability function for the adaptive display.

If affect can be classified with better than random rates, then potential users find themselves at a crossroads of sorts. When encountering technology with new capabilities users often find themselves in a “conceptual muddle” and administrators often find a “policy vacuum” [Moor, 1985].

In order to navigate through these muddles and vacuums, it is important to be aware of which human values are impacted by the introduction of new technology like displays that respond to affective state.

Value-Sensitive Design (VSD) [Friedman and Kahn, 2002] provides a value-ethics approach to the design of technology. It articulates a set of human values that might be impacted by new technology. Calmness, for instance is one human value that a soldier facing tough situations might desire. The value of autonomy is harder to place in the context of soldiering. Commanders are given a certain degree of autonomy in planning missions. If an adaptive display interferes with this through constant distraction, it may hinder a commander's ability to plan an execute a mission. Researchers developing adaptive displays that could be used by soldiers and commanders should consider the value Informed Consent. In the United States "research conducted by federal civilian employees or military personnel" is constrained by the Common Rule [Code of Federal Regulations, 1994], that requires informed consent of research participants.

The entire list of VSD values (some of which easily relate to adaptive displays) is: Human Welfare, Ownership and Property, Privacy, Freedom From Bias, Universal Usability, Trust, Autonomy, Informed Consent, Accountability, Identity, Calmness, and Environmental Sustainability. It may be difficult to think about how values like “Environmental Sustainability” in the most traditional sense relate to adaptive displays. However, the exercise of considering them is of use to administrators and individuals facing questions about adopting new technology. Let's now consider Calmness, Autonomy, and Informed Consent in more detail.

Calmness

One interesting project that relates to Calmness is Collective Calm [Reynolds, 2004]. This is a tug of war game in which teams compete to be the most calm. Here calmness is operationalized as a least squares fit of the slope of a skin conductivity graph. The project uses HandWave Bluetooth skin conductivity sensors [Reynolds, 2004]. These form piconets of physiology sensors that can communicate with displays like 3D games.

The importance of calmness has been considered particularly in the case of ubiquitous computing [Weiser and Brown, 1996]. If sensors and computers are to be pervasively embedded in the world around we don't want them making us agitated. Likewise, when soldiers cooperatively work together, calmness can be just as important.

The Tangible Media Group has considered various ambient displays that support calmness and other values in the context of cooperative work and architectural space [Wisneski et al., 1998]. This work relates to Value- Sensitive Design by providing working examples of how attributes like calmness can be supported through design.

Autonomy

Autonomy is another human value that is of relevance to developers of adaptive displays. Friedman and Nissenbaum discuss "system capability," "system complexity," "misrepresentation of the system,"and "system fluidity" as aspects of systems that can support or undermine autonomy [1997]. These aspects act on the individuals ability to "decide, plan, and act in ways that they believe will help them to achieve their goals" [Friedman and Kahn, 2002].

If a display provides fluidity to help users meet changing goals, then it may support autonomy. However, if the adaptive process that updates the display does not allow for user control [Hook, 1997] then the added complexity may undermine autonomy.

Still, without empirical data the relationship between an adaptive display and autonomy is speculative. What would be interesting is to see if adults report an adaptive display impacts their autonomy. One potentially interesting application that might be a good starting point for evaluation is IMWatching.

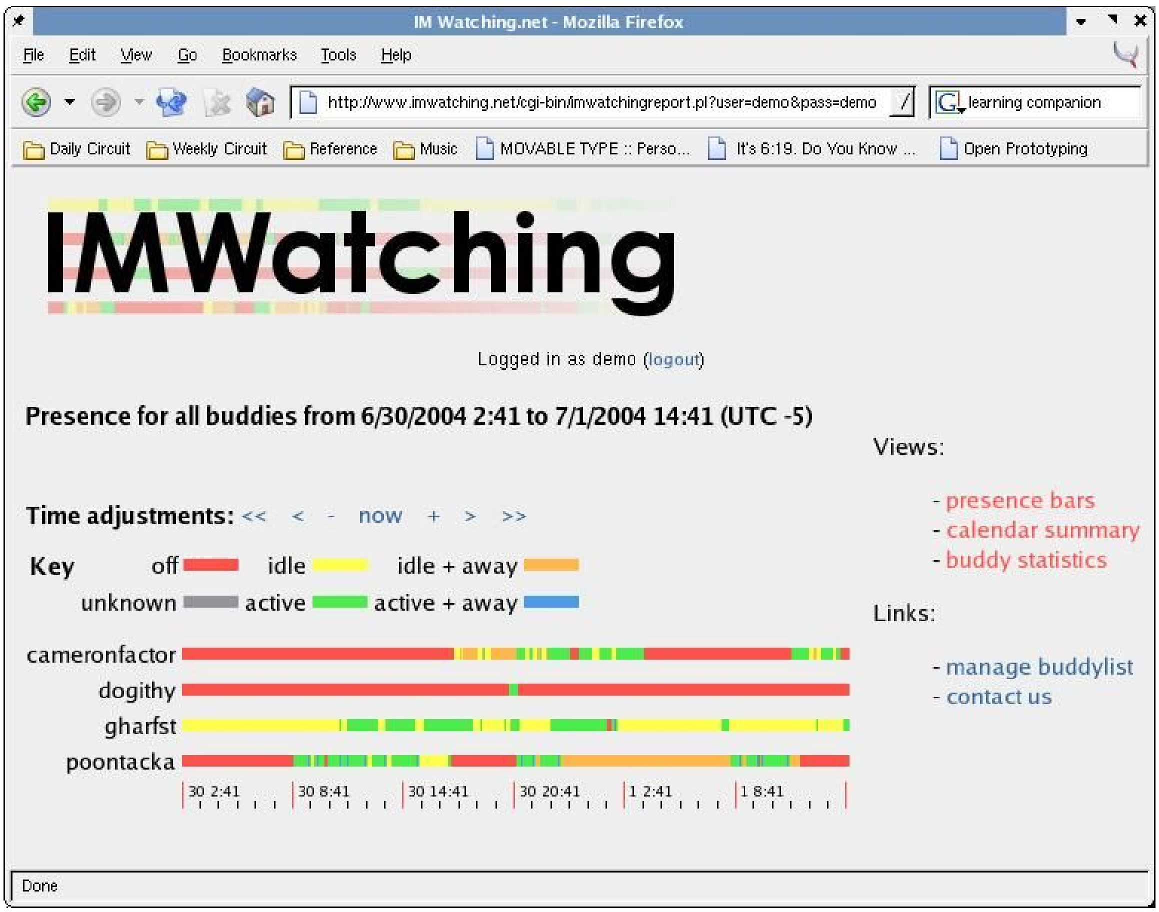

Illustration 1: IMWatching provides presence information without notifying that the instant messaging user is being observed.

IMWatching [Harfst, 2004] charts the instant messaging habits of a group of users. While instant messaging information is freely available and built into the AIM protocol and client, having long term displays of this activity spurs interesting ethical questions. A project that provides an interesting counterpoint to IMWatching is Watchme [Marmasse, 2004].

Watchme also provides presence information, but has an entirely different intent. The project focuses on closely- knit groups and furthermore requires reciprocity in access to data as well as interaction displays. More clearly: when someone with Wathchme is observing, the observed user knows, which is not the case with IMWatching.

Another interesting project that has implications on autonomy and adaptive displays is the Learning Companion project [Burleson et al, 2004]. The learning companion project consists of an agent that aids learners by being responsive to affect. A potential part of the learning companion's role is to provide performance information to educators and parents.

Suppose a soldier is using a program like the learning companion. This isn't a far- fetched supposition, considering the drill instructor in training software like America's Army. But if the drill instructor agent has the additional capability of sensing the soldier's emotions, this may help or hinder the soldier's ability to train. It could help by being aware of which duties make the soldier most uncomfortable . Alternatively, it could hinder training by undermining the soldier's confidence through feedback of negative emotional states.

The impact of the Learning Companion or the drill instructor in America's Army on autonomy remains unclear. As researchers, we can speculate about and anticipate these problems. But experimental data remains a crucial part of this process.

Informed Consent

Intimately related to autonomy is informed consent. Friedman writes "The idea of 'informed encompasses disclosure and comprehension. ... In turn, the idea of 'consent' encompasses voluntariness , competence, and agreement" [Friedman, and Kahn , 2002]. Given this sort of decomposition of informed consent, what is informed consent with respect to adaptive displays of affective sensors?

One way of supporting informed consent would be to provide easily readable contracts. A good example of these would be the Creative Commons license [Lessig, 2004]. This license is often presented to users in an abbreviated summary with icons.

Illustration 2: Icons make the creative commons license easy to read.

An interesting exercise would be to create icons for the values highlighted by value sensitive design, and then to embed them into a contract for displays of affective sensor data. This way, an individual confronted with a decision about use can easily read and consent.

We have now Calmness, Autonomy, and Informed Consent in more depth. At several points it has been suggested that decisions about the values that are impacted need empirical and experimental validation. What would such an experiment look like?

In Affective Sensors, Privacy, and Ethical Contracts we conducted just such an experiment [2004]. We hypothesized that ethical contracts improve ethical acceptability. Here ethical acceptability was gathered on questionnaires by asking about topics of “respect for privacy,” “willingness to use,” and “comfort.”

After two pilot studies were conducted we conducted the primary survey. This consisted of a balanced randomized 2x4 design. Between subjects, there was a control or contract condition. Within subjects were two application contexts (music and news recommendation) that focused on four emotions (joy, anger, sadness, and excitement). In the hypothetical scenarios, the emotion was observed by the application and used to adapt its behavior.

We had a total of 64 participants. The primary result was that in the contract condition respect for privacy improved ($p=0.0004$). There was also a trend toward improvement of comfort ($p=0.08$). These results were collected using on- line questionnaires of textual scenarios.

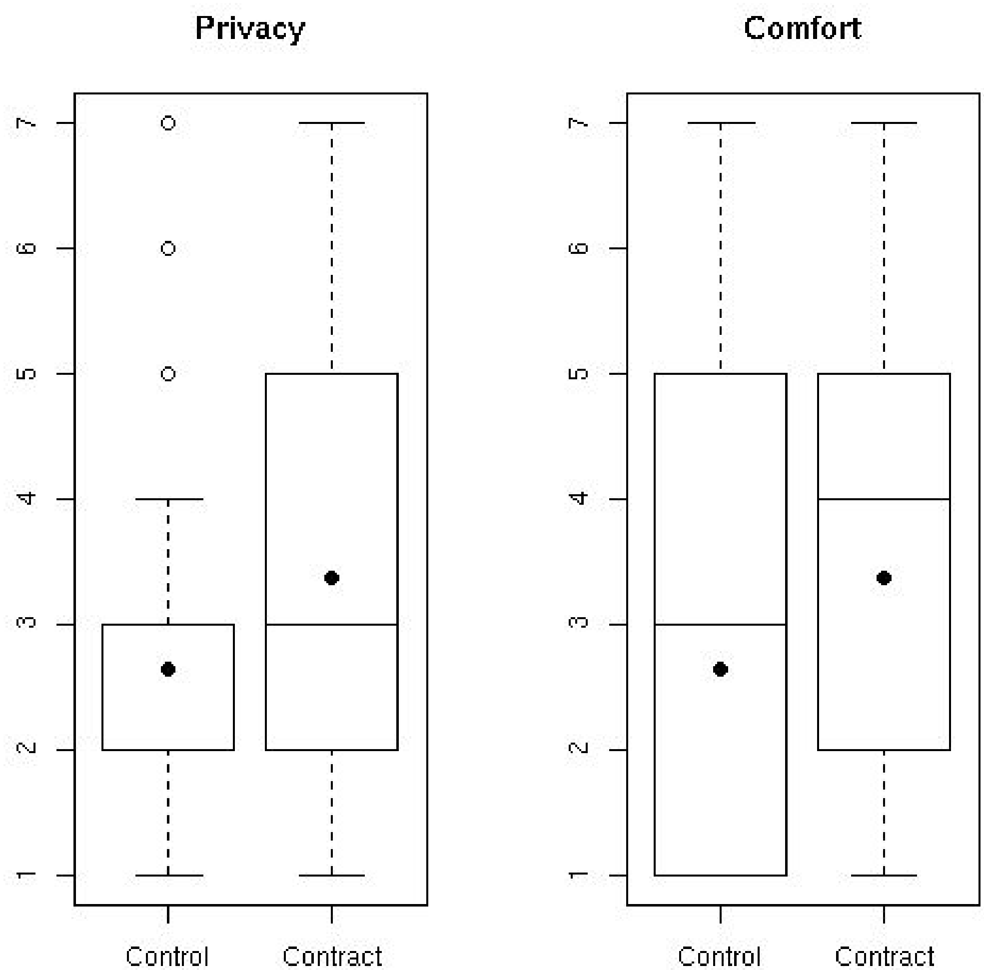

Illustration 3: Likert- scale data from ethical contract experiment. Participants reported a difference in privacy (p=0.0004), with a mode of 2 and mean of 2.6 in the control condition. This compares with a mode of 3 and a mean of 3.4 in the contract condition. The Privacy question was "Do you think your privacy would be affected ..." In the privacy boxplot (left) the scale is 1 - Completely Invaded to 7 - Completely Respected. A non-significant (p = 0.08) improvement in comfort was also observed. In this case, the question was "How comfortable would you feel ..." In the comfort boxplot (right) the scale is 1 - Completely Uncomfortable to 7 - Completely Comfortable.

Following this work, we are now considering a new experiment that will perform evaluations of live systems. A potential design for such an experiment might involve two participants communicating an emotional state with an applications like EmoteMail [Angesleva et al., 2004]. EmoteMail is an email client that communicates facial expressions and typing speed along with text. Following use, participants could be asked to fill out on-line questionnaires motivated by Value-Sensitive Design.

In describing the relationship between ethics and adaptive displays of affective sensor data, it is our intention to help individuals -- both users and those who make decisions about others' use -- encountering dilemmas. Hopefully, anticipating and reflecting on possible issues will help these individuals, be they researchers, administrators, or soldiers. Calmness, Autonomy, and Informed Consent are three issues that we hope are considered carefully and more throughly.

References

Angesleva, J., Reynolds, C., and O'Modhrain, S. (2004). Emotemail. (In Press) In Proceedings of the 31st International Conference on Computer Graphics and Interactive Techniques (SIGGRAPH 2004), August 8 -- August 12, Los Angeles, California.

Jameson, A. (2002). Adaptive interfaces and agents. In: The human-computer interaction handbook: fundamentals, evolving technologies and emerging applications, Pages 305--330.
http://www.cs.ubc.ca/~conati/532b/papers/handbook-chapter-jameson.pdf

Burleson, W., Picard, R. W., Perlin, L, and Lippincott, A. (2004). A platform for affective agent research. AAMAS workshop on empathetic agents, New York, USA.

Code of Federal Regulations (1994). Title 45 Section 46.116: General requirements for informed consent.

Department of Defense (1998). U.S. Marine Guidebook of Essential Subjects.

Department of the Army (2003). The Solder's Guide (FM 7-21.13 Section 1-91). Washington D.C., October 15, 2003.

Friedman, B. and Kahn, P. H., Jr. (2002) Human values, ethics, and design. In J. Jacko and A. Sears (eds.), Handbook of Human-Computer Interaction. Lawrence Erlbaum Associates, Mahwah NJ.

Friedman, B. and Nissenbaum, H. (1997). Software agents and user autonomy. In proceedings of the first international conference on autonomous agents. 466-469.

Peters, R. (2004). In praise of attrition. http://carlisle-www.army.mil/usawc/Parameters/04summer/peters.htm

Goldberg, D. (1989). Genetic algorithms in search, optimization, and machine learning. Reading, Massachusetts: Addison-Wesley Publishing Company.

Harfst, G. (2004). IMWatching.net: Record and Analyze Anybody's IM Presence. http://www.imwatching.net/

Hook, K. (1997).Evaluating the utility and usability of an adaptive hypermedia system. In Proceedings of the 2nd international conference on Intelligent user interfaces. 179--186. http://doi.acm.org/10.1145/238218.238320

Kapoor, A. and Picard, R.W. (2003). Fully Automatic Upper Facial Action Recognition. In Proceedings of Faces and Gestures (AMFG 2003), 17 October 2003, Nice, France

Kirpatrick S., Gelatt, C.D., Vecchi, M.P. (1983). Optimization by simulated annealing. *Science* 220, 671--680.

Lessig, L. (2004). Creative Commons Deed. http://creativecommons.org/licenses/by-nc/2.0/

Marmasse, N. (2004). Providing Lightweight Telepresence in Mobile Communication to Enhance Collaborative Living. Ph.D. Thesis. Massachusetts Institute of Technology , Cambridge , MA.

Moor, J.H. (1985). What is computer ethics? Metaphilosophy, 28(3) 266-275.

Picard, R.W. and Klein, J. (2002). Computers that Recognise and Respond to User Emotion: Theoretical and Practical Implications. Interacting with Computers, 14 (2) (2002), 141-169.

Qi, Y. and Picard, R.W. (2002). Context-sensitive Bayesian Classifiers and Application to Mouse Pressure Pattern Classification, in Proceedings of International Conference on Pattern Recognition, August 2002, Québec City, Canada.

Reynolds, C. (2001). The sensing and measurement of frustation with computers. Master's Thesis. Massachusetts Institute of Technology , Cambridge , MA.

Reynolds, C. (2002). A geometrical theory of adaptation. http://web.media.mit.edu/~carsonr/pdf/gtoa02.pdf

Reynolds, C. (2004). Collective Calm. http://mindgames.medialabeurope.org/collective_calm.htm

Reynolds, C. (2004). Cooking with Bluetooth http://arsenal.media.mit.edu/notebook/archives/000218.html

Reynolds, C. and Picard, R.W. (2004). "Affect Sensors, Privacy, and Ethical Contracts." Conference on Human Factors and Computing Systems (CHI 2004). Vienna: ACM, April 24-29, 2004.

Shewchuk, J.R. (1994). An Introduction to the Conjugate Gradient Method Without the Agonising Pain, Internal Report, School of Computer Science, Carnegie Mellon University, Pittsburgh, PA, USA, Aug. 1994 (available at http://www.cs.cmu.edu/~jrs/jrspapers.html).

USARIEM (1992). Technical Note 93-1. http://www.usariem.army.mil/somalia/operstre.htm

Weiser, M. and Brown, J.S. (1996). The Coming Age of Calm Technology. http://www.ubiq.com/hypertext/weiser/acmfuture2endnote.htm

Wisneski, C., Ishii, H., Dahley, A., Gorbet, M., Brave, S., Ullmer, B. and Yarin, P. (1998). Ambient Displays: Turning Architectual Space into an Interface between People and Digital Information, in Proceedings of International Workshop on Cooperative Buildings (CoBuild '98), (Darmstadt, Germany, February 1998), Springer Press, pp. 22-32.