Emily Long
elong1@umbc.edu

How Much Emotional Data Should Affective Computing Systems Record?

ABSTRACT

This paper reviews various literature written by affective computer researchers and theorists whose topics focus on ethics in affective computing, specifically on privacy and how much should affective computing compute. Following a review of the literature, this paper discusses the merits and pitfalls of the literature, provides a critique, and compares and contrasts the literature. It then concludes the findings of the literature reviewed and on the research question.

INTRODUCTION

As the technology of affective computing expands and grows, as will it's potential use expand and grow. Currently affective computing has a limited scope and breadth, with a good number of theories and hypothetical situations (Picard, 2003). Despite the newness of affective computing, researchers are exploring possible futures for affective computing, especially those fields related to security, safety, marketing, and decision making for a large group (Bullington, 2005). However, there are concerns about what emotional data should be examined by computers, and how that data is shared, stored, and processed, especially in a global environment where users' data is examined, processed, and sold for a company's benefit (Picard et al, 2004). These concerns are directly related to privacy, as well as how humans would eventually interact with affective computing in a negative way.

LITERATURE REVIEWED

This paper examines and observes a six papers that came from a variety of different years, author's, and perspectives, reflecting on a diversity within the research overall on affective computing and how much should affective computing compute. Two papers were written or co-written by Rosalind Picard, a leading researcher in the overall field of Affective Computing (Picard, 2003). Two papers also drew their inspiration from the science fiction TV show, *Black Mirror* (Cooney, et al 2005) (Mensio, et al 2018). All of the papers focus on the discussion on what should affective computing compute and the ethical implications involved in them.

Research in Affective Computing and Ethics

The paper "Beyond the basic emotions: what should affective computing compute?" focuses on the difference between basic and non-basic emotions, and what affective computing should compute beyond those basic emotions (De'Mello et al 2013). According to the paper, basic emotions include, angry, anxious, contempt, disgust, fearful, happy, sad, and surprised and the non-basic emotions as bored, confused, curious, delighted, eureka, engaged, and frustrated (De'Mello et al 2013). These distinctions between basic and non-basic emerged as a result of studies conducted in the 1960s (De'Mello et al 2013), at which point it is worth the

question as to whether or not these definitions of basic and non-basic are still relevant with today's knowledge and research.

The paper then analyzed five different studies that focused on emotions with academic tasks (De'Mello et al 2013). The studies involved 14,359 emotion reports, 131 participants, over 92 hours of interactions, and had a variety of different methods of collection, including emote-aloud, online self-reports, and cued-recall (De'Mello et al 2013). The analysis led the authors to gain three insights: first, that computer interaction are affectively charged; the second, it was more often that the neglected non-basic emotions were more prevalent than the basic emotions; and thirdly, engagement, boredom, confusion, and frustration were the most common emotions observed (De'Mello et al 2013).

The paper also discusses three limitations of this research (De'Mello et al 2013). The first being that they are drawing generalizations from a small number of studies, which could draw criticisms from the research community (De'Mello et al 2013). However, these claims that the researchers make were backed by statistically significant results (De'Mello et al 2013). The paper also talks about how all of these studies were only conducted using academic tasks, which is a major issue, as certain emotions might be more prevalent in academic settings that it would be in a different setting, such as a word setting (De'Mello et al 2013). For example, while workers and students might experience boredom and frustration, workers might appear more annoyed at their coworkers who they have to sit with all day, while students rotate in their classes who they sit next to and interact with. The third limitations discusses how the studies were conducted in a laboratory and thus might not be accurate when applied to the real world (De'Mello et al 2013).

While this paper is useful at looking at emotions and when they are appear at certain times, it doesn't really answer the question of "how much should affective computing compute?" It more so discusses the prevalence of basic and non-basic emotions. There is a brief discussion on affective computing and how it could observe basic and non-basic emotions, however that discussion does not delve further into the ethics or what should affective computing compute, which is ironic since it is the title of the paper. The paper also doesn't seem to question the validity of basic vs. non-basic emotions, or why non-basic emotions appear more often than the basic emotions. The studies mentioned in the paper should also be re-run in a variety of different settings in order to maintain their validity.

The first of two papers by Rosalind Picard, "Affective computing: challenges" explores the ethics within affective computing, with a particular focus on privacy (Picard et al, 2004). Picard introduces the idea that designers and users are in a contract, implied or real, in which designers create ethical design (Picard et al, 2004). Picard believes that designers should consider revealing the ethical decisions and thought process that they go through in order to be completely transparent with their users (Picard et al, 2004). Picard believes these basic ethical design considerations should be extended to affective computing systems: "who will collect emotional data, what type of emotions are recognized, and for what task the recognized emotional data is used" (Picard et al, 2004).

Picard and her team then conducted an evaluation “that examines privacy effects of affect sensors” (Picard et al 2004). The evaluation consisted of introducing users to a variety of scenarios that revolved around “emotions, sensors, and various contracts between designer and user” and then asked questions of the users through the form of a survey (Picard et al, 2004). Picard predicted that users would have a greater sense of privacy if they believed that their privacy was being invaded (Picard et al, 2004). The users responded that they would prefer that systems would have a greater respect for privacy, however the users did not have a problem with using a device if regardless of the respect for privacy (Picard et al, 2004).

While this paper is informative about user preferences into privacy, it still doesn’t necessarily address the question that the paper first brought up which was, “who will collect emotional data, what type of emotions are recognized, and for what task the recognized emotional data is used” (Picard et al 2004). This paper should address that question, even just theoretically, based on the results of the study mentioned. This paper should also address the questions of why users prefer a “respect for privacy” but they still would continue to use a system that may or may not have respect for their privacy.

Possible Affective Computing Futures

The paper “'Affective' computing and emotion recognition systems: the future of biometric surveillance?” discusses the future of biometric surveillance with affective computing (Bullington, 2005). The paper discusses three scenarios of how affective computing can be used in surveillance, primarily for security and safety concerns (Bullington, 2005). The first, and described as the most likely by the paper, is that of sensors and camera in the cockpit of a plane or a train where the pilot or driver could be observed in order to see if they are sleepy or distracted (Bullington, 2005). If the pilot/driver was in indeed sleepy, then the system would alert the crew or co-pilot/driver, which could allow them to take action if necessary (Bullington, 2005).

The second and not very likely scenario could be used for Group Decision making, using the affective system as a way to provide group feedback to the rest of the group (Bullington, 2005). This could also be used by decision makers or facilitators running a focus group in order to help them make an accurate decision (Bullington, 2005). This scenario does not seem like an effective method of using affective computing. Group members may not want their emotions projected to other members, especially if there is one member who disagrees with the rest.

The third and not very likely scenario would be using an affective system by financial institutions to determine if someone was committing fraud (Bullington, 2005). The paper doesn’t mention that an affective system could also be used by law enforcement institutions for surveillance of vulnerable areas or with interrogation techniques in order to determine if a suspect is lying, similar to a more accurate lie detector test. While the first scenario of monitoring a pilot is

very likely, the other scenarios are not very likely, and the law enforcement scenario would make more sense than the other systems.

These possible scenarios bring up privacy concerns, which are not fully explored in the paper. Firstly, all of these scenarios assume that the user would have to give consent to be observed by the system because the user would be aware of their emotions being measured and examined. That situation might work in the scenario of the pilot or law enforcement, but it could potentially be difficult for a grouping situation, such as a work group trying to make a decision on a direction that a project might go. If one member is feeling uneasy, they might not want the other group members to know their emotion if they feel like they are an outlier in the decision making.

Secondly, the paper does not address the actual means of security or what aspects of emotional data observed would be actually recorded. Would the plane system only observe and record tired or sleepy emotions, or would it also observe and record angry or frustrated emotions? In order to maintain safety of the individuals, emotional data would have be narrowed down so that they would only record the emotions that the system is specifically looking for.

Ethical Concerns of Affective Computing

The article, “Pitfalls of Affective Computing” focuses on various negatives of using affective computing in current society (Cooney, et al. 2005). It frequently mentions the popular science fiction show that focuses on our potential technological future, “Black Mirror,” as an inspiration for the paper (Cooney, et al. 2005). The paper focuses on four pitfalls of affective computing, psychological harm, physical harm, miscommunication, and disempowering individuals (Cooney, et al, 2005).

Psychological harm could involve revealing mental disorders of an individual, such as depression or anxiety, but it could also involve making it impossible to lie, even a white lie. (Cooney, et al 2005). The author’s argue that this would “force honesty”, and since everyone else would know your emotions, there would be no need for other forms of communication (Cooney, et al 2005). Physical harm could occur as the result of emotion bringing about sexual attacks or violence such as an individual seeing their partner have strong positive emotions with another person (Cooney, et al 2005) However, this could also lead to a positive or neutral effect in that a person could detect negative intentions using affective technology from a potential attacker and remove themselves from a potentially dangerous situation (Cooney, et al 2005).

The third pitfall is miscommunication, which would involve systems having to deal with the wide complexity of emotions, since most people feel a variety of emotions as one time (Cooney, et al 2005). This could result in the systems saying that someone is feeling one thing, when in reality they are feeling something completely different, which is people relied on the systems instead of their own intuition, this could result in miscommunications between individuals (Cooney, et al 2005).

The fourth and final pitfall would be disempowering individuals, which the authors believe would come about because computers and robots would use the emotional data as a way to manipulate humans into "liking" them or doing what they want (Cooney, et al 2005). However, the paper does not mention that other humans could use the technology in order to manipulate others, which might be a greater risk, since humans are more likely to manipulate that computers or robots. The paper then goes into ways to avoid these pitfalls, which involve using an off- button so that people could hide their emotions if they feel like they need to, also allowing for false emotions to be read so that individuals could keep their privacy of emotions (Cooney, et al 2005).

The paper "The Rise of Emotion-aware Conversational Agents" also drives it inspiration from *Black Mirror*, and specifically focuses on an episode where a main character became emotionally attached to a conversational agent of her deceased partner (Mensio, et al 2018). The paper primarily focuses on the development and then the negatives of a conversational agent that would be able to detect an individual's emotion and then modify the conversation to accommodate the individual's mood (Mensio, et al 2018). The paper breaks down the development into three different stages of "advancement" (Mensio, et al 2018).

The first is textual interactions, which is where a person would be communicating using conversation with an agent through a textual format (Mensio, et al 2018). A likely modern day example of this would be a Help Chat Bot that users could ask for basic help (Mensio, et al 2018). The second stage is vocal interactions, which is where a person would be communicating with an agent using verbal conversation (Mensio, et al 2018). So basically imagine Siri can understand your emotions when you're yelling at her for misinterpreting your speech. The third stage is embodied agents, which is where the conversational agent that would be programmed and designed to look like an actual human (Mensio, et al 2018). While a present day example of Sophia, a humanlike robot that can give pre-programmed answers depending on the question, but a science fiction example could be Data from *Star Trek: The Next Generation*, an android who looks and acts human and can react depending on the emotions of his crewmates (Mensio, et al 2018).

The paper then goes through potential threats to humanity through using these conversational agents (Mensio, et al 2018). The first is that the conversational agents really only provide a short-term therapeutic effect (Mensio, et al 2018). In the episode of *Black Mirror*, the characters only receive the benefit of interacting with the conversations initially, when they are opening up to the conversational agent as a means of comfort (Mensio, et al 2018). While this is a positive, the paper determines that this is just the beginning and can lead to a variety of other threats, including addiction, isolation, and a change of personality (Mensio, et al 2018). Addiction can come into play when the conversational agent is someone that the individual really relies on or potential misses, like in the episode of *Black Mirror*, the main character became reliant on the conversational agent that embodied her deceased partner for comfort (Mensio, et al 2018). This

addiction can lead to isolation, because the user becomes so reliant and find normal human conversations lacking that they isolate themselves from other humans (Mensio, et al 2018). This combination of addiction and isolation can result in individuals suffering a change in personality (Mensio, et al 2018).

It was interesting to find two papers that drive their inspiration from the show, *Black Mirror*, however, neither of these papers actually mentioned the specific episodes that drew their attention (Cooney, et al 2005) (Mensio, et al 2018). This little detail, while might seem inconsequential, makes it difficult for a reader to go and watch the particular episode mentioned and then compare their thoughts with the authors' opinions, thus creating a greater dialogue. The first paper also seems to just mention the show as an inspiration, but does not draw parallels again after the introduction, while the second paper continually makes references to the particular episode and refer back to it when discussing their ideas (Cooney, et al 2005) (Mensio, et al 2018). This allows the reader to understand the inspiration more, and if they have seen that particular episode mentioned, can make parallels and relate to the author more because there is a shared experience.

The second of two papers by Rosalind Picard, "Affective sensors, privacy, and ethical contracts" focuses on the challenges and criticisms for affective sensors, privacy, and ethical concerns and then counteracts those challenges and criticisms with her own arguments (Picard, 2003). The criticisms a broken up into separate groups, with the first two being that many modalities are not accessible by affective technology (blood chemistry, brain activity, etc) and the people's expression is so individual and varied that it is nearly impossible to determine various emotions (Picard, 2003). Picard argues that there will be a need for computers to read emotions like humans can and have been able to for eons (Picard, 2003). Picard also countered that her own studies had found that a skin surface sensor could indeed determine emotion of the wearer 80% of the time over the course of two weeks (Picard, 2003).The third criticism is that there has been limited with cognitive modelling, which would assist in developing effective computer emotion modeling (Picard, 2003). Picard argues that the models themselves are likely flawed, in that they rely on stereotypical personalities and exaggerated traits (Picard, 2003).The fourth criticism is that because computers do not have a physical body, they cannot reliably express believe emotion (Picard, 2003). Picard argues that computers can indeed express emotion through a variety of different methods besides just their physical bodies, such as signals or cute noises, like R2D2 of Star Wars (Picard, 2003).

Systems can also be designed to embody the appropriate emotion based on the emotion expressed (Picard, 2003). The paper then discusses more about the ethics of computers reading emotions, in that emotions are highly personal and private and thus should not be detected by computers, but Picard argues that our perception of emotions is not private because humans themselves routinely detect, recognize, and redirect emotions of other people (Picard, 2003). It seems strange

that we expect computers to act like humans and yet we also put restriction so that they don't act like humans.

While I agree with Picard that we can expect computers to act like humans in that they can read emotions, there is also the thought that what if computers could read emotions better than other humans. Someone who has trained themselves to be an expert at hiding their emotions might suddenly find their cool head disrupted by a computer being able to read their emotion even if their fellow humans might not. At the very beginning of this paper, Picard mention the famous *Star Trek* character, Mr. Spock, a half-Vulcan who is known for keeping his emotions in check, making him seem like he is lacking in emotions (Picard, 2003). Picard argues that while Mr. Spock may not outwardly express his emotions, he is still experiencing a wide variety of emotions while keeping his emotions hidden from his captain and crewmates (Picard, 2003). This insight is similar to the point made in the "Pitfalls of Affective Computing" paper, in that it would require humans to have a sense of false honesty when dealing with someone who is interacting with an affective device, making it difficult for humans to keep their emotions private (Cooney, et al 2005).

DISCUSSION

After reviewing the literature individually, it is interesting to examine them as a wider group and how they compare to each other. On the higher level of two of these papers being written by Rosalind Picard, and that two of these papers drew their inspiration from *Black Mirror*, these papers all seek to answer or address the overarching question of what are ethical practices in Affective Computing when it comes to what emotional data should be observed and record. The papers do not seem to have a conclusive answer to this subject. One of the papers, discussed a that users do want to have some sort of a respect for privacy, but would still continue to use a system that did not respect their privacy (Picard et al, 2004). Another of the papers brought up the question of what kinds of emotions should be observed by affective computing systems, basic vs. non-basic, but that paper focused specifically on the presence of nob-basic emotions (De'Mello et al 2013). Several papers addressed the potential negatives and warned about potential issues on society with affective computing. Some of the authors' logic makes sense, specifically with the paper that focuses on conversational agents (Mensio, et al 2018). If humans were to have a means of companionship that could act perfectly depending on their emotions ever single time, then it begs the question of why have interactions with other humans at all (Mensio, et al 2018).

Originally, this author was more interested in ethics related to commercial benefit, however, this paper has evolved into focusing more on the greater ethical concerns of the usage of affective computing. If we intend for computers to act just like humans in their emotional needs (Picard 2003), then based on the research there could be grave consequences (Mensio, et al 2018). It is wise for researchers to limit their development in affective technologies to uses where the type of

emotion is limited and the reasoning is helpful, such as observing a pilot or a driver who might become sleepy, and thus a danger to themselves and others (Bullington, 2005). These affective devices would require complete consent from the user and would benefit more than just the user and the system itself. Because the users of these devices would not find pleasure in these interactions, the users would not also become addicted to the device. Overall, those types of affective devices would improve society and not compromise user data at the same time.

CONCLUSION

After reviewing and analyzing the literature, it is apparent that there are pitfalls and issues that are coming with the advancement of affective computing technology. As affective computing technology improves over time, as the need to insure that the privacy of the technology is kept secure, and to insure that the relationship between humans and computers is kept positive. It is vital that designers and researchers consider these aspects when they are developing affective computing technology. This gives designers and researchers a duty to insure that their affective computing technology does compromise user data and does not hamper the lives of the individuals who use those devices.

CITED SOURCES

1. Bullington, J. (2005). 'Affective' computing and emotion recognition systems: the future of biometric surveillance? *InfoSecCD '05: Proceedings of the 2nd annual conference on Information security curriculum development.* September 2005. 95-99. https://dl.acm.org/doi/10.1145/1107622.1107644
2. Cooney, M., Pashami, S., Sant'Anna, A., Fan, Y., & Nowazck, S. (2018). Pitfalls of Affective Computing: How can the automatic visual communication of emotions lead to harm, and what can be done to mitigate such risks. *WWW '18: Companion Proceedings of the The Web Conference 2018.* April 2018. 1563-1566. https://dl.acm.org/doi/10.1145/3184558.3191611
3. De'Mello, S. & Calvo, R. (2013). Beyond the basic emotions: what should affective computing compute? *CHI EA '13: CHI '13 Extended Abstracts on Human Factors in Computing Systems.* April 2013. 2287-2294. https://dl.acm.org/doi/10.1145/2468356.2468751
4. Mensio, M., Rizzo, G., & Morisio, M. (2018). The Rise of Emotion-aware Conversational Agents: Threats in Digital Emotions. *WWW '18: Companion Proceedings of the The Web Conference 2018.* April 2018. 1541-1544. https://dl.acm.org/doi/10.1145/3184558.3191607
5. Picard, R. W. (2003). Affective computing: challenges. *International Journal of Human-Computer Studies, 59(1-2).* 55-64. https://www.sciencedirect.com/science/article/pii/S1071581903000521
6. Picard, R. & Carson, R. (2004). Affective sensors, privacy, and ethical contracts. CHI EA '04: CHI '04 Extended Abstracts on Human Factors in

Computing Systems. April 2004. 1103-1106.
https://dl.acm.org/doi/10.1145/985921.985999