CHAPTER THREE

Boiling the frog: Ethical leniency due to prior exposure to technology

Noah Ari[a], **Nusrath Jahan**[a], **Johnathan Mell**[a], **and Pamela Wisniewski**[b]
[a]University of Central Florida, Orlando, FL, United States
[b]Vanderbilt University, Nashville, TN, United States

3.1. Introduction

To create effective and meaningful metrics for ethical human-centered artificial intelligence (AI) systems, we must understand how and why people make the choices they do. If we do not, design frameworks will not reflect human decision making and consequently will lack ecological validity. When considering how to architect AI systems with these constraints, we must understand the ethical and practical decisions people make in real-world situations. This will enable us to design software that can understand those decisions in context. In addition to understanding real-world decision making, we also must consider how to evaluate what people care about while balancing ethical concerns. How often does a person make a minor unethical decision for larger ethical benefit later? How can AI be robust enough to avoid ethical confusion by this concept while also behaving in a way that corresponds with how humans want it to behave?

The impact AI systems and technology in general have on us is also an interesting factor when considering the design of social virtual agents. To design agents that can effectively interact with humans requires these agents to navigate the confusing heterogeneity and fungibility of human ethos. To explore the challenges that human ethics present to agent design, we explored emotion detection in the context of a simple economic game. Past work shows people feel that emotions are personal and sensitive and betray their intentions and actions [4]. Other work also shows ethical evaluation is subject to a habituation effect in the domain of negotiation [7]. To create agents that interface well with humans, emotional intelligence concerns must be considered [6].

Trolley Crash. https://doi.org/10.1016/B978-0-44-315991-6.00009-1

People tend to evaluate emotions as personal and sensitive information, which provide insight into behavior in such a way that creates a sense of vulnerability [4]. The breakneck pace of emotion detection software development is at odds with our understanding of the ethical and social impact of increased access to emotional data.

Questions on the balance between self-interest and ethical mores are critical to advancing the field of human-like agents. Yet they remain difficult to answer. There are limited experimental data forming the grounds about how people behave when presented with AI technology and how the presence of this technology affects human decision making and the capacity for ethical evaluation of that technology. The crux of our study is understanding how people use controversial technology. This chapter will focus on the rapidly developing domain of emotion detection that is unfortunately coincident with a lack of empirical research into developing ethical frameworks for their use [1]. Our study is designed to contribute empirical research to help advance the data-driven pursuit of responsible emotion detection.

To study the impact of pre-exposure to a technology on the behavior of the pre-exposed user and the ethical evaluation of that technology by the pre-exposed user, we have formulated the following research questions (RQs):

RQ1: Does prior exposure to technology make participants view that technology more favorably?

RQ2: Does prior exposure to emotion detection technology make it more likely that participants will choose to use it?

RQ3: Does the publicity of one's usage of systems like emotion detection change people's likelihood of choosing to use it in the first place?

RQ4: Does the presence of emotion detection alter the behavioral patterns of participants, especially their decisions within a competitive economic game?

Open-ended RQ: Do people find emotional detection software to be invasive or unethical?

3.2. Background

The disconnect between technological advancement and society's acceptance of it is known as "cultural lag." Cultural lag [5] is a well-known phenomenon that occurs during periods where material culture, in this case technological development of emotion detection, outpaces non-material

culture such as our ethical frameworks, infrastructure, and laws surrounding that technology. When considering cultural lag in combination with past research that indicates ethical decision making can be habituated toward less ethical decisions in competitive contexts like negotiation [7], it becomes clear that exposure to technology could create risks to proper ethical framework formulation. Recent reviews have found participants disliked emotion recognition and considered it "intrusive" and even "scary" [4], yet these technologies are appearing more and more in our society [1]. This collides with our understandings of ethical leniency and cultural lag and raises the following questions: Does exposure to emotional detection technology alter our perceptions and behaviors? Would such a change be detectable by the one experiencing it?

3.3. Literature review

3.3.1 The use of emotion detection in online contexts

Due to its immense usefulness, there are numerous domains developing and using emotion detection technology in the real world [1]. Video surveillance to detect human behavior has propitious use in public safety and security applications [3]. Emotion detection technology is being used increasingly in various online contexts, such as capturing emotions in multimedia tagging [2]. Visual data or multimedia data that contain several emotions from humans have been used for detecting emotion or human behavior. Methods for mining information about the behavior of humans have been developed and improved [3]. Moreover, there are surveys for investigating technologies capable of emotion detection. The strengths and shortcomings of those technologies have been identified, pointing out the areas of emotion detection technology where further research is required [1]. The literature review demonstrated the lack of research on emotion detection technology from an ethical point of view. It provided an intuition into the ethical aspects of emotion detection to be researched as previous technologies analyzed the effect of emotions in our bodies, omitting the behavioral impact. Furthermore, these studies explored the factors that contributed to rationalization as a likely explanation for the rapid spread of technologies and the prolific use of algorithms that violate privacy [8]. However, more empirical explorations of these principles are necessary. We contribute to the growing corpus of ethical examinations in the specific field of emotion detection as opposed to generalized privacy.

3.3.2 The ethical considerations of emotion detection

Emotions are rated by people to be intimate. Data on human emotions are sensitive and rich in information that provides insights into behavior [4]. People also believe that evaluating a person's emotional state is subjective and raises a host of ethical questions [9]. Andalibi and Buss found several consistent themes regarding both properties of emotional data and perceptions of emotional recognition. The nature of emotional data corresponds to their findings in the similar theme that participants disliked emotion recognition and regarded it as "intrusive" and even "scary" [4,9]. Andalibi and Buss also noted that participants regularly mentioned the difficulty people have evaluating the emotions of others and even themselves. This difficulty went on to inform a stated thematic uneasiness among participants regarding predictive uses of emotion detection. Overall, this intrusiveness and uneasiness combine to form a larger theme of general distrust of the practice of emotion recognition [4]. Consequently, researchers focused extensively on the intersection of emotional data and AI ethics considering the social effect of the emotion detection or recognition technologies [9]. It has been shown that risky technologies are essentially required to be analyzed from the point of ethical acceptability along with social acceptance [6], because mere social acceptance studies are not capable of adequately catching all the moral attributes of risky technologies. Several studies [4,6,9] found answers to questions that ground ethical concerns about risky technologies with particular focus on ethical aspects of emotional recognition technology's ascent toward ubiquity online [4]. The participants resoundingly echoed the theme that emotional data are often extremely sensitive and personal, and the applications necessitating the collection of such data had varying rates of approval. While these papers explore what people think of risky emotion detection and recognition technologies in an interview format regarding their potential and current uses, we intended to explore what people do when presented with one such technology and whether personal use alters their general ethical evaluation of this type of technology.

3.3.3 Technology acceptance and habituation

In human–computer interaction (HCI) research where interest is concentrated on technology-related human and social effects, investigating and exploring different aspects of technology adoption is appropriate and required [10]. The potential ethical issues resulting from the widespread rollout of any technology need to be researched and considered before its

introduction [6]. Prior research has shown that people judge negotiation strategies as more ethical if they have prior experience with negotiations [7]. Mell et al. conducted a study to find out how negotiating experience influences human endorsement of negotiation tactics and strategies. They showed that people adopt more deceptive and manipulative tactics if they have prior negative experience [7]. This study found that participants who interacted with tough agents were more willing to endorse negotiation tactics that involved deception and manipulation. These results demonstrate that ethical evaluations are subject to a habituation effect in the context of negotiation tactics. In our task, we examined a similar question in the context of the evaluation of emotion detection technology. We hypothesized that prior exposure to emotion detection technology would render the attitude of people towards using this technology more positive, evaluating it as a more ethical technology than those who do not have prior experience.

3.3.4 Evaluation of technology

Past research has defined two main properties that determine the acceptance of technology. Those properties are perceived ease of use and perceived usefulness. Past work created scale items to evaluate these properties [11]. The technology acceptance model (TAM) serves as a framework for understanding the likelihood of someone adopting a given technology. We used the TAM to inform our design. Since the TAM considers these two general properties as influencing adoption, we designed our system to be as useful and participant-friendly as possible to prevent people from opting not to use the system for design reasons over ethical ones. This helped ensure the core ethical dilemma remained the focus of this experiment. We chose to use technology adoption propensity (TAP) [12] for actively measuring adoption propensity since for the case of this experiment, we were more interested in the participants' beliefs about technology and emotion detection than how easy systems like emotion detection are to use. People's TAP can accurately be measured using a 14-item questionnaire consisting of two contributing (optimism and proficiency) and two inhibiting (dependence and vulnerability) factors. TAP can predict people's technology use behaviors [12]. The following hypotheses (Hs) were developed to quantify the problem and proposed solutions to be investigated as part of the experiment.

3.4. Problem

H1: Participants who use emotion detection in the first game will rate emotion detection as more ethical than those who did not.
H2: Participants who use emotion detection in the first game will be more likely to choose to use it again in the second game.
H3: Participants who are told their opponent will know their decision to use or not use emotion detection in the second game will be less likely to choose to use emotion detection in that game.
H4: Those with access to emotion detection will be more likely to choose to steal than those without it.

3.5. Methods

We conducted a 2 × 2 between-subjects experimental design. The experiment was divided into two game phases and three surveys. The condition for the phase 1 game and subsequent survey was the randomized exposure to the emotion detection system. The phase 2 game and final survey had an additional user-selected factor taking the place of the randomized condition in the prior phase; however, this user selection was also connected to the phase 2 condition which was the framing of the self-selection. The phase 2 condition was whether the decision to self-select the use of emotion detection in the second game was framed such that the participant's opponent would be informed or would not be informed of the participant's decision regarding the use of the technology. Fig. 3.1 shows a visual representation of the full process a participant underwent. The experimental design consists of participants playing a short game against a virtual agent presented as a human opponent and answering a short survey. Over the course of the experiment, one participant went through each of the following steps. The participant first answered a survey that tested ethical perceptions of emotion detection technology. Afterward, the participant was randomly assigned to the emotion detection group or the group without emotion detection and was then asked to complete a 10-round iterated prisoner's dilemma game called the split–steal game. After playing the 10-round game, the participants were asked to complete the same survey to measure their ethical perception of emotion detection.

Upon completion of the second iteration of the ethics survey construct, participants were introduced to the emotion detection technology if they did not use it in the first game. All players regardless of their use of emotion

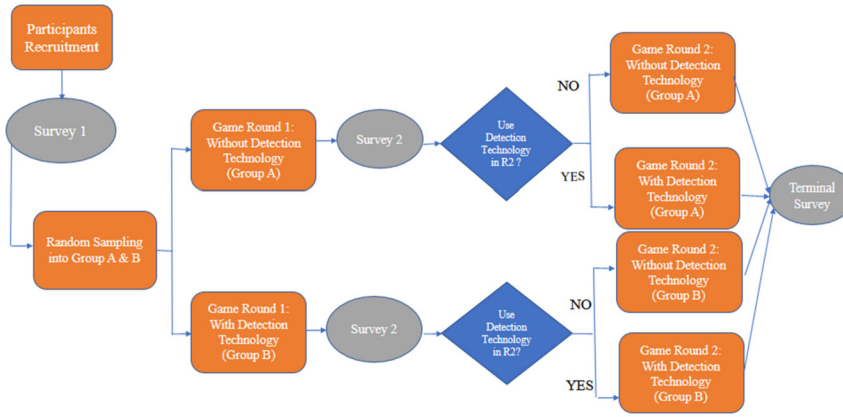

Figure 3.1 A diagram of the experimental design.

detection were then provided the opportunity to use emotion detection in a second iteration of the game given that their opponent would not be provided this opportunity and would have no knowledge of the participant's usage of emotion detection. After making their choice, they played the game again with or without emotion detection. After the second game participants were asked to complete the survey once more, measuring their ethical evaluation of emotion detection.

The game architecture utilized is an economic game called the split–steal game, adapted from the architecture presented by Hoegan et al. [13] as an iterated 10-round prisoner's dilemma game. Each participant played against a virtual agent and was told that the virtual agent was another human player. This deception framed the experiment [14] in such a way as to avoid the differences in the ways that humans treat virtual agents as opposed to humans [13,15] to maintain the strength of the ethical framing of the experiment. The virtual agent was not represented as an avatar and users were provided with a false loading screen to support the illusion of human–human gameplay. Similarly, prior to the experiment participants were instructed they may have to turn on their camera but were always subsequently informed that they were placed into a group that does not require a camera and told "Therefore you will not be asked to turn on your camera." In the first instance of playing this game, participants were randomly granted access or not to a simulated emotion detection technology as an aid against their virtual opponent. In the second instance, participants were asked whether they wanted to use the emotion detection technology

or not. There were minor changes of wording depending on the first condition ("Would you like to use the technology?" vs. "Would you like to use the technology again?"). Each participant played 10 rounds consecutively against the same agent. All virtual agents followed a tit-for-tat strategy aside from the first two rounds where the agent first cooperated (split) and then defected (steal). This behavior was chosen as a neutral point to examine effects separate from strategic considerations as demonstrated in existing literature [13].

The survey used prevalidated measures [16–18] to assess the participants' evaluation of general emotion detection software in terms of how ethical they believe it to be and how useful they believe it to be. In addition, attention checks were added into the surveys to maintain high standards of attention for online participants.

3.5.1 Measures

Our main independent factor was prior exposure of the participant to the emotion detection technology. This was operationalized by the two conditions where the technology was introduced at different points in the process.

We examined the effects of the use of emotion detection in the second game on participants' decisions, the participants' answers to the seven-point Likert scale survey questions regarding ethicality and favorability, and finally the participants' answers to the open-ended survey questions at the end. In addition to the self-reported Likert measures and main ethical decisions, we examined the game actions of the participants, including their pattern of "split" vs. "steal" actions.

3.6. Data analysis

3.6.1 Ethical leniency (H1)

Quantitative results. After testing and concluding there was no failure of random assignment and our constructs were reliable, we conducted a repeated measures analysis of variance (ANOVA) on the dependent variable of the ethical evaluation construct averages against the independent factor of pre-exposure (group assignment). Fig. 3.2 visualizes the survey responses of participants where the p column value is an indication of whether or not the ANOVA measurement is statistically significant. We found significance ($p = 0.023$) in the construct of ethicality over time, but no significance

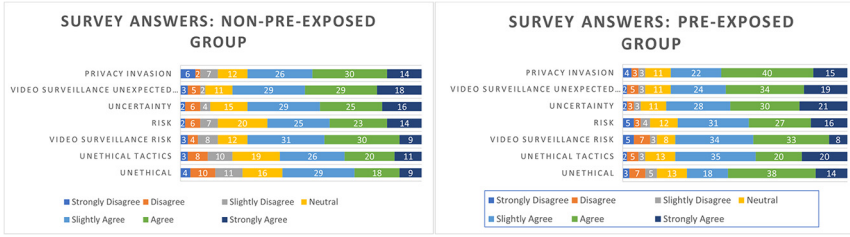

Figure 3.2 Stacked bar graphs demonstrating the correlation between pre-exposure and rating emotion detection as more ethical or less unethical.

against the independent factor ($p = 0.122$) in contradiction to the results found in our initial pilot. In the initial study of $n = 56$, where n represents the number of participants, we found a significance of $p = 0.007$. This in conjunction with the significance we found isolated from the independent factor prompted some follow-up analysis utilizing paired t-tests. When running paired t-tests on the construct averages, the pre-exposed group had a strong significance ($p = 0.008$) while the non-exposed group had no significance ($p = 0.28$). These results were interesting, because they implied that there was a significant time effect driven exclusively by the pre-exposed group. This suggests that a habituation effect is present but is not exclusively determined by pre-exposure but is possibly the result of contributory effects. The Likert scale questions are coded such that "strongly agree" refers to agreeing that technology is unethical. For instance, strongly agreeing to "privacy invasion" means strongly agreeing that emotion detection risks invasions of privacy.

Qualitative results. Every participant who was pre-exposed to emotion detection voiced that they specifically felt comfortable with the technology because of the unfair advantage it gave them in the game. Some participants even went so far as to state without prompting that they were very comfortable with the emotion detection but would not have been comfortable if it was directed at them. One participant stated that "Since it wasn't directed at me, I had no concerns" regarding their comfort with the technology, while another stated that their "comfort level was fine, as long as it is not my emotion being detected."

3.6.2 Likelihood of adoption (H2)

Quantitative results. To evaluate results for H2, we conducted a Pearson chi-square test on the dependent variable of choice to use emotion detection in the second game against the factor of the participants' random

group assignment to be pre-exposed or not to determine significance. We found marginal significance ($p = 0.51$), indicating an increased likelihood of choosing to use the technology among the pre-exposed group. This was expected from the pilot study, where initially there was no significance since some conditional groups were too sparsely populated for robust and conclusive analysis. With the increased n of 168, the initial trend remains the same with 129 participants choosing to use the technology and only 39 choosing not to, implying that increasing n could further solidify the finding of a significant effect.

Qualitative results. The thematic analysis conducted on the initial pilot study's results informed the framing of this study to address emergent themes of curiosity, triviality, and self-serving behavior, as shown in Table 3.1, which dictates limitations of the initial study.

3.6.3 Known usage

Quantitative results. To evaluate H3, we conducted a Pearson chi-square test on the dependent variable of the decision to use the emotion detection tool in the second game against the independent factor of the framing group the participants were in. We found a strong significance of $p = 0.036$, indicating a significant effect on the choice to use the technology when both players would know if the participant chose to use it, showing that participants were far less likely to use the technology when their opponent would know of their decision to use it.

3.6.4 Behavioral effects

Quantitative results. To evaluate H4, we conducted a Pearson chi-square test against the dependent variable of the frequency of the selection of the steal option per participant against the independent factor of use of emotion detection during the game. We found a strong significance of $p < 0.001$, indicating a strong modification of game behavior based on the presence of the emotion prediction software. Fig. 3.3 shows these behavior patterns. As the figure demonstrates, the significance is in the direction of higher steal counts becoming significantly more common for those using the emotion detection tool.

Qualitative results. In our analysis of the open-ended question regarding whether people believe their behavior changed or would change depending on the presence of emotion detection, we found a very interesting dichotomy. There was a divide amongst the participants between believing

Table 3.1 A collection of the results of the thematic analysis conducted on the initial study group.

Theme	Code		Exemplars
Perception of behavioral effect	Believes they played differently than they would have played because of emotion detection ($n = 17$)	Stole more due to use of emotion detection ($n = 6$)	• I would have probably **not chosen to steal so much**. • Yes **I'd have split a lot more** on choices. • If I didn't use the emotion detection, **I might have used split more, but I probably would have lost points overall**.
		Split more due to use of emotion detection ($n = 1$)	• **I think I would have. I would have stolen on the first turn if I didn't use the emotion detection**. I would do this because I would not trust the other player to do what is best for both players. But since I knew what my opponent was likely to do, **I was swayed to splitting based on the emotion detection results**.
	Believes they did not play differently ($n = 26$)		• No, I don't think so. **I always like to be fair and impartial, so my playing style would remain the same**, regardless. • Probably not, **I think the expected outcome is best if you steal every time**.

continued on next page

Table 3.1 (*continued*)

Theme	Code	Exemplars
Decision justification	Selfishness ($n = 13$)	• **Since it wasn't directed at me I had no concerns**. • In the game I was okay if no faces or voices are captured and shown to others. **If my face or voice were needed to be captured, I'd not take the hit**. I think knowing their emotion made it easy for me to decide to split or steal. • It was alright, it made me more successful in my decision-making about whether to steal or not. **My comfort level was fine, as long as it's not my emotions being detected**.
	Curiosity ($n = 6$)	• I wanted to **try something new**. • [I] chose to use it (...) to **see how well it would work**.
	Triviality ($n = 7$)	• **Because it was a game, I felt like it was not that big of a deal to use it**. I felt like, for this experiment, I was comfortable using it because it was just a game. • I felt it was just a game. **If it was something important like a real-life business negotiation, I would feel uncomfortable and unethical doing it**.

continued on next page

Table 3.1 (*continued*)

Theme	Code	Exemplars
Comfort level	Comfortable ($n = 30$)	• **Emotion detection made me feel comfortable in this game**. I don't have to think or [sic] afraid of [the] opponent's choice. • **It was quitely [sic] comfortable because I [sic]** afraid what the other person chose.
	Uncomfortable ($n = 3$)	• **It made me feel like I was taking advantage of the other player**. I was uncomfortable knowing too much about the other player and felt like it gave me an **unfair advantage**. • **Very uncomfortable**. My opponent was unaware of the emotion detection activities and, to me, **it was a form of cheating**. • **I was uncomfortable with the presence of emotion detection in the game because I felt I knew extra information that should have been private**.

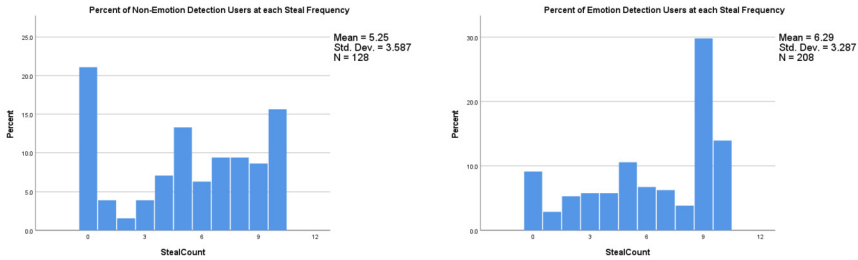

Figure 3.3 Participants' frequency of the steal decision for both participants using the emotion detection tool (left) and participants not using the system (right).

their behavior was or would be impacted depending on the presence of emotion detection in the game and the non-insignificant tendency toward participants evaluating their game behavior as unaffected by the presence of the emotion detection technology. After our thematic analysis, we found that $n = 17$ participants believed that their behavior did change or would have changed, while $n = 25$ participants believed that their behavior did not change or would not have changed. Despite acknowledgment that the emotion detection technology made the game easier, participants often still stated that they believed they played largely the same way they would have without the emotion detection. One participant remarked that the emotion detection "made their opponent easier to understand" but also said that they believed they "wouldn't have played different [sic], but I [the participant] would have been slower at the game" because they would take more time to make the decision. These results are quite interesting given the strong significance for behavioral alteration found in the quantitative results and suggests a more interesting effect than simply utilizing a tool during gameplay as the sole reason or implication for this behavioral shift.

3.7. Use cases

While there are immediate uses to adding novel technologies into society, for example using tools or agents in a detective capacity, such tools may impact the behavior of the humans using them. In this case, it is worth noting that evaluating the impact of technology on behavior may be necessary before deploying certain technologies in sensitive or precise contexts. When developing systems, researchers will also have to consider that this effect could influence their judgment as well, demonstrating the need for evaluative research and diverse opinions on novel technologies. Further-

more, when designing robots and virtual agents who are meant to interact with people on a regular basis, the mere presence of the agent using observant technologies like emotion detection can impact the behavior and judgment of the humans around it.

3.8. Applications

Applying the findings of this paper is difficult in the short term (see future work) but our study provides quantitative and qualitative evidence that considerations must be made when designing socially interactive or observant systems that coexist with human users. Not all behavioral patterns are guaranteed to be the same across domains and this must also be considered. Though human behavior or reactions may be expected based on past data and thorough empirical research, the very presence of these novel systems and tools may modify the perceptions and behaviors of those around them.

3.9. Discussion

The main aim of our study was to examine the potential habituation effects of emotion detection technology on participants' evaluation of that technology, their decision to use that technology, and even their behavior in and outside of game. Below, we discuss our results with direct reference to the original research questions and by synthesizing our quantitative and qualitative results.

3.9.1 Ethical evaluation

Upon conducting a repeated measures ANOVA controlling for the condition (ethical evaluation by participants), we found there was a main effect of time ($p = 0.023$). Yet, the interaction term was not significant, presenting only a marginal trend. After conducting follow-up paired *t*-tests on the pre-exposed and non-exposed groups we found that this time effect was driven primarily by the pre-exposure, since the paired *t*-test for that group showed significance ($p = 0.008$), whereas the non-exposed group did not ($p = 0.280$). Therefore, we hypothesize that habituation does indeed play a key role. However, the mere process of asking ethical questions in a repeated fashion seems to also have a significant effect. What we found suggests that both this "wearing down" of participants by asking them ethical questions in a repeated fashion and the habituation effect exist as contributory effects

and are difficult to disentangle. Researchers who are regularly exposed to the technologies that they develop themselves or humans who use these technologies regularly are likely subject to the same habituation effect. This means that the capacity of a researcher to evaluate their own technology as ethical or unethical is dampened by their proximity to it, not just due to any expected conflict of interest, but also due to this habituation effect. This also goes for research that involves asking participants to use a technology like emotion detection and then asking them to evaluate it as ethical or unethical, as well as research that investigates perceptions of technology without consideration of who has or has not used that technology in the past.

3.9.2 Adoption

Like the initial study, the group who chose not to use the emotion detection was quite small. Only $n = 39$ participants chose not to use the technology while $n = 129$ chose to use the technology. The marginal significance found implies that there may exist a main effect of pre-exposure to a technology and the choice to use it again, which we hypothesize would become apparent by increasing the number of participants further. In conjunction with the qualitative data, this suggests that a habituation effect may lead people to compromise their ethical stance on a given technology and choose to use it anyway over time.

3.9.3 Publicity of usage

These results provide insight into how the visibility of one's decision to adopt or use a technology interacts with the habituation effects found. A strong main effect was found on the decision to use emotion detection when the participant was told their opponent would be informed of their decision. This suggests that though participants are more likely to use a technology based on the habituation effect found in H2, this effect could be offset or mitigated through its conflict with the effect found in H3.

3.9.4 Behavior

These results show that pre-exposure to the emotion detection system presented had a direct and strong effect on the behavior of participants. This is particularly interesting because of the direction of the effect, showing that participants who used the technology are far more likely to choose to "steal" than to "split" in the game they were presented. When considering

the formulation of the iterated prisoner's dilemma, this behavior is already interesting since the best outcome is obtained when both participants split each time. The agent, as a tit-for-tat player, would have been just as likely to split as to steal dependent upon the participant's decision. Steal, however, is often the most immediately apparent option since it had the most expected return over the course of one individual round. Ultimately, this shows further behavioral impact of the presence of this emotion detection technology on participants.

3.10. Conclusions

The ethical evaluation of emotion detection and the effects at play in evaluating such technologies are important to understand, especially when considering the design and implementation of AI and agent-based systems in human-centered and human-interactive environments. We believe that to properly evaluate the rapidly rising domain of emotion detection as well as considering how to interface agents with humans ethically and effectively, we must better understand the impacts of interacting with novel technologies like emotion detection on our perception of the technology and behavior while using the technology. Similarly, to construct virtual agents that effectively navigate human interaction these agents will have to be constructed keeping in mind the effect that they have on humans. This includes the moral implication of the virtual agent's existence in human spaces and well as behavioral impacts explored in this study. Our results demonstrate the habituation effect we expected, though due to the lack of significance against pre-exposure on its own, the results suggest that the habituation effect is connected to contributory effects from time and pre-exposure. Furthermore, our results demonstrated behavioral impacts on those using the technology that went largely undetected by the participants. We contribute our work as an extension of work surrounding emotion detection, ethical computing, and agent design.

3.11. Outlook and future works

The study was limited by only being able to accommodate play against one agent strategy. This does not negatively impact the quality of the results in the sense that regardless of agent behavior, the agent was constant across the two groups. However, there is a possibility that different agent strategies or behaviors could have unique impacts on participant behavior

as well, especially in a context where the agents are presented as humans and the iterated game is explicitly conducted against the same opponent. Mechanisms such as trust or reputation could also have an impact on behavioral results. Future work should examine the effect of pre-exposure against varying agent behaviors to examine the role of trust and reputation with the opponent on the pre-exposure effect discussed in this paper. Future work may also examine the behavioral impact of emotion detection or unfair technological advantages. Future work may vary the framing of the access to the emotion detection as unilateral, bilateral, or individual. This study exclusively focused on the unilateral use of emotion detection to mitigate any strategic concerns to best isolate the ethical component as the primary focus of this experiment. However, variations that state that a choice to use emotion detection grants it to your opponent as well or that your opponent is given the same choice may provide interesting insights into variations of this effect.

Notes and acknowledgments

Portions of this work previously appeared at the International Conference on Intelligent Virtual Agents (IVA'22). New results have been included in this book chapter.

References

[1] J.M. Garcia-Garcia, V.M. Penichet, M.D. Lozano, Emotion detection: a technology review, in: Proceedings of the XVIII International Conference on Human Computer Interaction, 2017.
[2] S. Wang, et al., Capturing emotion distribution for multimedia emotion tagging, IEEE Transactions on Affective Computing (2019).
[3] T. Ko, A survey on behavior analysis in video surveillance for homeland security applications, in: 2008 37th IEEE Applied Imagery Pattern Recognition Workshop, IEEE, 2008.
[4] N. Andalibi, J. Buss, The human in emotion recognition on social media: Attitudes, outcomes, risks, in: Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems, 2020.
[5] R.L. Brinkman, J.E. Brinkman, Cultural lag: Conception and theory, International Journal of Social Economics (1997).
[6] B. Taebi, Bridging the gap between social acceptance and ethical acceptability, Risk Analysis 37 (10) (2017) 1817–1827.
[7] J. Mell, et al., The effects of experience on deception in human-agent negotiation, Journal of Artificial Intelligence Research 68 (2020) 633–660.
[8] N.J. Fast, A.S. Jago, Privacy matters… Or does it? Algorithms, rationalization, and the erosion of concern for privacy, Current Opinion in Psychology 31 (2020) 44–48.
[9] A. McStay, P. Pavliscak, Emotional Artificial Intelligence: Guidelines for Ethical Use, EmotionalAI.org, 2019.
[10] J. Lindley, P. Coulton, M. Sturdee, Implications for adoption, in: Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems, 2017.
[11] F.D. Davis, Perceived usefulness, perceived ease of use, and user acceptance of information technology, MIS Quarterly 13 (3) (September 1989) 318–340.

[12] M. Ratchford, M. Barnhart, Development and validation of the technology adoption propensity (TAP) index, Journal of Business Research 65 (8) (2012) 1209–1215.
[13] R. Hoegen, et al., Comparing behavior towards humans and virtual humans in a social dilemma, in: International Conference on Intelligent Virtual Agents, Springer, 2015.
[14] A. Tversky, D. Kahneman, Multiple Criteria Decision Making and Risk Analysis Using Microcomputers, Springer, Berlin, Heidelberg, 1989.
[15] B. Reeves, C. Nass, The Media Equation: How People Treat Computers, Television, and New Media Like Real People and Places, Cambridge University Press, 1996.
[16] N.K. Malhotra, S.S. Kim, J. Agarwal, Internet users' information privacy concerns (IUIPC): The construct, the scale, and a causal model, Information Systems Research 15 (4) (2004) 336–355.
[17] N. Michaelidou, M. Micevski, Consumers' ethical perceptions of social media analytics practices: Risks, benefits and potential outcomes, Journal of Business Research 104 (2019) 576–586.
[18] S.L. Jarvenpaa, N. Tractinsky, L. Saarinen, Consumer trust in an Internet store: A cross-cultural validation, Journal of Computer-Mediated Communication 5 (2) (1999) JCMC526.