The Covert Control of Affective Computing: Psychological Manipulation Effects of Emotion-Based Personalization

Xinyu Leng

Abstract:

The deep integration of big data and the platform economy has triggered a fundamental shift in the application paradigm of affective computing, transforming it from a service-oriented approach to a control-oriented one. We critically examine the operational mechanisms and societal harms of affective computing as a covert control tool. First, we dissect its technological foundation: capturing emotions through multimodal data to quantify sentiment, then enabling predictive interventions into user psychology via dynamic modeling and affective matching algorithms. Subsequently, we systematically elucidate four core psychological manipulation mechanisms rooted in affective personalization: emotional contagion and atmosphere engineering, cognitive narrowing and information silos, operant conditioning and behavioral shaping, and decision hijacking and impulse exploitation. The research reveals that these psychological manipulation mechanisms constitute a comprehensive and profound erosion of individual autonomy, the foundations of democratic society, commercial ethics, and personal privacy and dignity. This paper calls for the establishment of an appropriate ethical and legal regulatory framework to harness technological risks and defend human agency in the digital age.

Keywords: Affective computing; Psychological manipulation; Personalization; Algorithmic ethics

1. Introduction

The theoretical goal of affective computing is to endow machines with the ability to recognize, understand, process, and even simulate human emotions. This aims to bridge the persistent emotional gap in human-computer interaction and build a more natu-ral, harmonious, and intelligent human-computer environment. In its early technological vision, affective computing was envisioned for application in fields with significant social benefits, such as autism thera-py assistance, fatigue driving monitoring, and intelligent education [1]. By enabling empathetic responses to human psychological states through computational

models, it aimed to enhance the efficiency and well-being of human-machine collaboration. This phase primarily focused on emotion signal acquisition and pattern recognition algorithm optimization, embodying a distinct blend of instrumental rationality and humanistic care. However, with the maturation and widespread adoption of technologies like big data, deep learning, and cloud computing, the application scenarios of affective computing underwent fundamental expansion and transformation. When deeply integrated with commercial capital and platform economies, emotional computing has been widely deployed in precision marketing, content recommendation, social interaction, and digital entertainment. This deployment has quietly shifted its underlying logic from a service-oriented paradigm to a control-oriented one [2]. Consequently, the core objective of the technology is no longer limited to passively responding to user needs to optimize experiences. Instead, it has evolved into actively and proactively intervening in users' psychological processes, shifting from understanding emotions to guiding behaviors toward commercial or political objectives. For instance, e-commerce platforms analyze users' micro-expressions and browsing behaviors in real time to infer emotional states, then push product information designed to trigger impulse purchases. News aggregation algorithms capture users' emotional resonance points to continuously amplify specific content types, prolonging user engagement or reinforcing ideological positions. This transforms affective computing from a neutral technological tool into a power mechanism embedded within the capillaries of digital life.

Personalized technologies based on emotional states have become highly sophisticated psychological manipulation tools. The covert nature of their operation exponentially amplifies their potential social risks and ethical harms. Its "precision" surpasses traditional user profiling based on historical behavior, directly targeting the underlying drivers of human decision-making to achieve real-time, dynamic modeling and intervention of individual psychological dynamics. Its 'stealth' stems from mechanisms deeply embedded within seamless user experiences, disguised as "thoughtful service" and "personalized care," rendering manipulation imperceptible to conscious awareness. These characteristics render it more pervasive and effective than any traditional form of persuasion or advertising, while simultaneously making it harder to regulate and hold accountable. Consequently, this paper will delve into the technological foundations enabling affective computing to achieve "covert control," focusing specifically on the mechanisms of its psychological manipulation effects. It will analyze how processes such as emotional contagion, cognitive narrowing, and operant conditioning are realized within personalized interactions. Furthermore, this paper

will explore the multifaceted harms and profound ethical dilemmas that technological evolution poses to individual autonomy, social structures, and democratic processes. Building upon this analysis, it calls for the establishment of an appropriate governance framework. The nature of power revealed under affective computing provides a critical theoretical perspective for understanding and regulating emerging technological forces.

2. Relevant Theory and Technical Foundations

2.1 Data Capture of Emotions

The foundation of affective computing relies on multimodal data acquisition technologies to fuse and transform humans' intrinsic, unstructured emotional states into machine-readable, computable quantitative data. This has established a relatively mature paradigm in both academic research and industrial practice. Existing literature consistently indicates that single data sources have limitations in emotion recognition, making multimodal fusion the mainstream approach in this field. Specifically, multimodal fusion approaches include: First, facial expression recognition. Based on Ekman's cross-cultural theory of basic emotions, analyzing facial action units (AUs) through computer vision has become a key indicator for emotion recognition [3]. However, the distinction between social smiles and genuine joy makes this method susceptible to deception. Second, speech sentiment analysis. Extracting acoustic features from speech signals—such as fundamental frequency, energy, and speech rate—enables inference of the speaker's emotional state. For instance, anger and joy exhibit distinguishable acoustic patterns, though background noise and cultural variations remain practical challenges. Third, physiological signal monitoring. Physiological signals like heart rate variability (HRV), galvanic skin response (GSR), and electroencephalogram (EEG) are often regarded as more "authentic" emotion indicators due to their lower degree of conscious control. However, the invasive nature of collecting such data limits their widespread daily application [4]. Fourth is text and behavioral data analysis. In natural language processing (NLP), sentiment analysis determines the emotional polarity of text through vocabulary, syntax, and contextual cues. Simultaneously, user behavioral data on digital interfaces serves as a proxy variable for inferring interests and emotional engagement.

Despite significant advances in emotion digitization, academic circles maintain profound critiques of its theoretical foundations. Simplifying complex, fluid, and embodied

emotional experiences into discrete, decontextualized data points constitutes a form of "computational reductionism." This approach inevitably loses the rich personal meaning and sociocultural connotations inherent in emotions. The digitization process replaces authentic human emotional experiences with machine-operable "data emotions."

2.2 Building Personalized Models

After completing emotion digitization, research focuses shift to leveraging this data to construct dynamic personalized models for precise prediction and intervention. First, the user emotional profile—a high-dimensional, dynamic psychological model updated through continuous data streams—constitutes the core output of this process. Sequence models based on Recurrent Neural Networks (RNN) or Long Short-Term Memory (LSTM) networks can predict a user's most probable emotional response tendency at specific times and under specific content stimuli, based on their historical emotional data. This enables platforms to transcend understanding users as "who they were in the past" and instead predict "how they will feel next." Second, algorithmic recommendation and emotional matching represent the model's application layer [5]. Existing literature reveals two primary manipulation strategies: First, emotional resonance. Systems that push content aligned with a user's current emotional state enhance immersion and stickiness. Emotionally consistent content is more likely to be shared and endorsed, thereby reinforcing information silos. Second, emotional regulation. Systems steer users' emotional states toward desired directions [6]. When detecting low mood, platforms may push comforting product ads (resonance); when sensing anger, they might deliver news sparking righteous indignation to boost engagement (regulation). The system becomes an "emotional curator," whose logic prioritizes maximizing platform engagement over user well-being.

The academic discourse surrounding personalized models—transforming data into actionable predictive engines—is fraught with tension. On one hand, the technology's precision is astonishing; on the other, its application raises profound ethical concerns. Through predictive modeling, systems not only understand us but can preemptively shape our emotional responses, thereby subtly stripping individuals of their autonomy to experience and reflect upon their own feelings.

2.3 The Hidden Origins of Control's "Covertness"

The effectiveness and danger of emotional manipulation stem fundamentally from the covertness of its operational mechanisms. Relevant literature primarily identifies two levels of origin. On one hand, user-friendly interfaces and optimized experiences form the sugar coating of manipulation. Positive user experiences are instrumentalized to conceal manipulative intent. When manipulation manifests as thoughtful suggestions from smart assistants or personalized recommendations, users' psychological defenses struggle to activate. Thus, technology's social acceptability becomes the perfect cover for its power dynamics. On the other hand, subliminal stimuli and implicit guidance form the psychological core of manipulation [7]. Guidance in digital environments no longer relies on fleeting, subconscious flashes of imagery but is achieved through a comprehensive language of interaction design. For instance, choice architecture and default options can influence user decisions; micro-interaction designs like like notifications and red dot prompts leverage dopamine-driven intermittent reward mechanisms to unconsciously condition users into dependency on platforms. The characteristic of embedding control deeply within seamless experiences and cognitive processes transforms emotion-based manipulation into a form of governance—exercising power by shaping desires and steering emotions. As users navigate these services, their emotional autonomy is quietly eroded, often without their awareness.

3. Psychological Mechanisms Underlying Emotion Personalization

The manipulation achieved by affective computing systems through emotion personalization technology relies not on overt coercion but on a series of deep-seated psychological mechanisms. We will delve into four core psychological mechanisms: emotional contagion and atmosphere creation, cognitive narrowing and information silos, operant conditioning and behavior shaping, and decision hijacking and impulse exploitation.

3.1 Emotional Contagion and Atmosphere Cultivation

Emotional contagion originally referred to the unconscious process by which individuals unconsciously capture and feel others' emotions through automatic mimicry and feedback. In digital environments, emotional contagion is systematically mediated and scaled by algorithms. By analyzing users' real-time emotional data, affective computing systems can precisely cultivate and sustain a specific informational atmosphere—for example, continuously pushing large volumes of success-oriented content to manufacture illusory optimism [8]. The psychological efficacy of algorithm-driven atmosphere creation lies in its persistence and personalization. Systems construct highly

emotionally consistent information environments, immersing users long-term in specific emotional tones that subsequently influence individual information processing styles and judgments. Users in anxious states tend toward heuristic information processing, becoming more receptive to fear-related framing; conversely, users in pleasant states exhibit relatively weakened critical thinking. Thus, the carefully orchestrated emotional contagion engineered by algorithms profoundly shapes users' overall mindset and fundamental perceptions of the real world, creating an algorithmically modulated emotional reality.

3.2 Cognitive Narrowing and the Information Bubble

In catering to user preferences, personalized emotional systems inevitably lead to cognitive narrowing. Content with high emotional arousal inherently captures and monopolizes attention resources more readily. To maximize user engagement, algorithms prioritize recommending such emotionally charged content, creating a positive feedback loop of "emotion-attention-recommendation." Prolonged exposure to this cycle confines users within information silos and echo chambers constructed by specific emotional tones [9]. Within these information silos, divergent viewpoints and facts are systematically filtered out. Users encounter information that continuously reinforces and validates their existing emotional biases and perspectives. This leads to a persistent narrowing of cognitive horizons, making it difficult to engage with or comprehend complex information that contradicts pre-established emotional stances. Furthermore, critical reflection and dialectical analysis—essential for rational thinking—gradually deteriorate due to the absence of challenging perspectives. Ultimately, users' cognitive structures become fragile and extreme, with their worldviews no longer grounded in careful evaluation of diverse facts but built upon singular emotional narratives amplified by algorithms.

3.3 Operant Conditioning and Behavioral Shaping

Emotional computing platforms essentially function as massive operant conditioning apparatus. Reinforcement or punishment of behavioral outcomes determines the probability of that behavior recurring. On platforms like social media, social positive feedback mechanisms—such as "likes," "shares," and "comments"—serve as potent variable-ratio reinforcers, forming reinforcement patterns that solidify and addictive behaviors. When a user's specific emotional expression receives unexpectedly high social rewards, they unconsciously adjust future expression strategies to sustain this social validation and

pleasure [10]. They increasingly adopt emotional patterns favored by algorithms and communities—a process of algorithm-guided behavioral shaping. This can subtly alter users' online personas and even offline emotional response patterns, sacrificing the authenticity and autonomy of their emotional expression for adaptive compliance with the platform's reward mechanisms.

3.4 Decision Hijacking and Impulse Exploitation

During emotional fluctuations, humans' rational cognitive systems become suppressed, relying instead on emotional, intuitive heuristic processing. Emotion computing systems, through real-time sentiment recognition, can precisely identify users' emotional vulnerability windows. Systems exploit these windows for so-called decision hijacking [11]. For instance, when users seek solace amid sadness, they receive targeted ads for high-margin "comfort consumption" like luxury goods or sugary treats. Or when users feel righteous indignation over an event, they are pushed extreme political views or petition links. During these moments of diminished rational judgment, users become significantly more susceptible to making irrational consumption decisions or readily accepting extreme viewpoints. Decision hijacking and impulse exploitation mechanisms fundamentally commercialize and politicize human cognitive vulnerabilities. They deprive individuals of the opportunity to make more prudent decisions once their emotions have settled, instead converting fleeting emotional impulses directly into immediate behavioral outputs.

4. The Hazards and Ethical Dilemmas of Affective Computing

4.1 Erosion of Individual Autonomy

Autonomy refers to an individual's capacity for self-governance based on rational reflection and personal values. However, emotion-driven personalized systems are pre-emptively steering decision-making pathways by shaping internal emotional and cognitive environments. When systems can accurately predict and exploit users' emotional vulnerabilities to push specific content or products, the critical thinking process—relying on careful reflection and information weighing—is bypassed. Unbeknownst to individuals, their decisions increasingly stem from algorithmically orchestrated emotional stimuli rather than their own rational deliberation and value judgments. This effectively reduces users to algorithmically guided "emotional puppets," with their intellectual freedom and behavioral

autonomy eroded invisibly. Emotional computing's erosion of autonomy is covert and cumulative—it does not negate choice but reshapes the chooser themselves, violating the fundamental principle that "human beings are ends in themselves, not merely means."

4.2 Threats to Society's Healthy Functioning

Society's healthy functioning relies on citizens possessing rational critical thinking and a public sphere enabling open, rational dialogue on public issues. The personalized application of affective computing poses a fundamental threat to this. First, at the level of political manipulation, affective computing technology enables micro-targeting to an extreme degree. Political actors can analyze collective emotional data of voters to design and deploy highly personalized propaganda messages that trigger specific emotional responses. Irrational emotional drivers bypass citizens' critical reflection, undermining the public rationality foundation of democratic decision-making by stoking intergroup antagonism and exploiting cognitive biases. Second, at the level of social cohesion, personalized emotional information flows exacerbate societal fragmentation. When algorithms persistently push content with opposing emotional tones and divergent factual frameworks to groups holding differing positions, they reinforce group polarization and the cognitive isolation effect. A shared sense of reality is replaced by mutually exclusive, emotionally opposed versions of reality. This not only hinders mutual understanding between groups but also makes it easier to view opponents as moral enemies rather than political rivals. Consequently, social cooperation mechanisms based on compromise and consensus become unsustainable, undermining the very foundations of society.

4.3 Ethical Deterioration in Business

In the commercial sphere, the application of affective computing distorts the traditional "customer-first" principle into systematic exploitation of users' psychological vulnerabilities, resulting in severe ethical lapses. By analyzing emotional data, companies transform users' states of anxiety, insecurity, and social cravings into targetable marketing opportunities, creating a form of capitalistic exploitation of emotional fragility. Emotional computing practices in commerce fuel and intensify excessive consumerism, encouraging users to address emotional needs through consumption rather than seeking genuine, non-commodified solutions [12]. Simultaneously, addictive design—leveraging dopamine-driven intermittent reward mechanisms—is widely adopted to deliberately cultivate psychological dependence on platforms or products. Fur-

thermore, practices such as exploiting health anxieties to sell unnecessary insurance or leveraging appearance insecurities to promote cosmetic products constitute unfair trade practices. By exploiting asymmetries in information and cognitive power to access users' most private emotional domains, these tactics violate the principles of fairness, integrity, and non-harm in commercial transactions.

4.4 Infringement of Privacy and Dignity

Emotional data is not ordinary behavioral data; it is among the most core and sensitive biometric information, directly reflecting an individual's inner state, personality traits, and value concerns. The continuous collection and deep analysis of emotional data constitute an invasion of one's inner world. Traditional privacy concepts primarily shield individuals from intrusions into external behavior and physical space. Emotion computing, however, targets deeper psychological privacy or the integrity of the mind [13]. When companies monitor, quantify, and exploit individuals' emotions for profit without explicit, informed consent, they effectively commodify humanity's most intimate emotional experiences into exploitable data resources. This severely infringes upon human dignity by denying individuals their status as subjects possessing emotional depth and self-determination, reducing them to predictable, manipulable objects. Consequently, safeguarding emotional data transcends mere privacy concerns—it becomes a core issue in preserving the fundamental autonomy and integrity that defines humanity.

5. Conclusion

We systematically demonstrate how affective computing has diverged from its original humanistic vision, evolving into a sophisticated and covert form of social control. Research indicates that emotion-based personalization is not merely a technical optimization but a novel paradigm of power operation. By quantifying humanity's most intimate emotional experiences, affective computing constructs dynamic psychological models. It then exploits deep-seated psychological mechanisms—such as emotional contagion, cognitive narrowing, conditioned reflexes, and decision hijacking—to systematically shape users' emotions, cognition, and behavior without their awareness. The consequences of affective computing are profound and multidimensional: at the micro level, it erodes individuals' critical thinking and autonomous decision-making abilities, reducing people to emotional puppets; On a macro level, it poisons public discourse through political micro-targeting and emotional information flows, exacerbating social divisions and undermining democratic foundations. In commerce, it exploits emotional vulnerability as a tool,

fueling excessive consumerism and addictive design. Ultimately, the collection and analysis of emotional data constitute the most fundamental violation of psychological privacy and human dignity.

Facing the severe challenges posed by affective computing, simply rejecting the technology is not viable. We must establish a robust ethical and governance framework to guide its responsible development. On one hand, at the level of technological ethics and legal regulation, emotional data must be designated as the most sensitive category requiring special protection, with legislation restricting its collection and usage scope. Algorithmic transparency and the principle of informed intervention must be implemented, granting users the right to know whether and how their emotional data is used for personalized recommendations, along with the right to refuse covert manipulation. On the other hand, tech companies must embed ethical design into their product development processes. They should provide pure information models free from emotional manipulation or allow users to disable emotion-based analytics. Furthermore, at the levels of public literacy and societal discourse, there is an urgent need to enhance the digital emotional intelligence and critical thinking skills of the entire population. Through education and social awareness campaigns, the public must be empowered to clearly recognize the mechanisms of emotional manipulation in the digital environment, cultivate metacognitive abilities regarding their own emotional states, and thereby build necessary psychological defenses.

The trajectory of affective computing is by no means predetermined. It can either become a tool that exacerbates control and inequality or serve the noble goal of enhancing human well-being and social intelligence. Through profound critique, prudent regulation, and proactive construction, we must ensure that the steering wheel of technological development remains firmly in the hands of human values.

References

[1] Bessi, A., Zollo, F., Del Vicario, M., Scala, A., Caldarelli, G., & Quattrociocchi, W. (2015). Trend of narratives in the age of misinformation. PloS one, 10(8), e0134641.

[2] Calvo, R. A., & D'Mello, S. (2010). Affect detection: An interdisciplinary review of models, methods, and their applications. IEEE Transactions on affective computing, 1(1), 18-37.

[3] Grind, S. (2018). How to build habit-forming products. Retrieved from YouTube. com: https://www. youtube. com/watch.

[4] Hildebrandt, M. (2015). Smart technologies and the end (s) of law: novel entanglements of law and technology. In Smart Technologies and the End (s) of Law. Edward Elgar Publishing.

[5] Hildebrandt, M., & O'Hara, K. (2020). Introduction: Life and the law in the era of data-driven agency. In Life and the law in the era of data-driven agency (pp. 1-15). Edward Elgar Publishing.

[6] Ho, A. G., & Siu, K. W. M. G. (2012). Emotion design, emotional design, emotionalize design: A review on their relationships from a new perspective. The Design Journal, 15(1), 9-32.

[7] Pariser, E. (2011). The filter bubble: What the Internet is hiding from you. penguin UK.

[8] Pariser, E. (2011). The filter bubble: How the new personalized web is changing what we read and how we think. Penguin.

[9] Nagulendra, S., & Vassileva, J. (2014, September). Understanding and controlling the filter bubble through interactive visualization: a user study. In Proceedings of the 25th ACM conference on Hypertext and social media (pp. 107-115).

[10] Zuboff, S. (2019). The age of surveillance capitalism: The fight for a human future at the new frontier of power, edn. PublicAffairs, New York.

[11] Haggart, B. (2019). The age of surveillance capitalism: The fight for a human future at the new frontier of power, S. Zuboff (2018). journal of digital media & policy, 10(2), 229-243.

[12] Zuboff, S. (2024). The age of surveillance capitalism: The fight for a human future at the new frontier of power. Journal of Information Ethics, 33(1), 84-85.

[13] Whitcomb, C. G. (2020). Review of Shoshana Zuboff (2019). The Age of Surveillance Capitalism: The Fight for a Human Future at the New Frontier of Power: New York: Public Affairs. 704 pp. ISBN 9781781256848 (Hardcover). Postdigital Science and Education, 2(2), 484-488.