Machine In The Middle: Exploring Dark Patterns of Emotional Human-Computer Integration Through Media Art

Rod Dickinson
The University of the West of England
Bristol, UK
rod.dickinson@uwe.ac.uk

Nathan Semertzidis
Exertion Games Lab, Human Centred-Computing, Monash University
Clayton, Australia
nathan@exertiongameslab.org

Florian 'Floyd' Mueller
Exertion Games Lab, Human Centred-Computing, Monash University
Clayton, Australia
floyd@exertiongameslab.org

ABSTRACT

As our relationship with machines becomes evermore intimate, we observe increasing efforts in the quantification of human emotion, which has historically generated unintended consequences. We acknowledge an amplification of this trend through recent technological developments that aim toward human-computer integration, and explore the dark patterns that may arise for integrating emotions with machinic processes through "machine_in_the_middle". Machine_in_the_middle is an interactive system in which a participant wears an electroencephalographic headset, and their neural activity is analysed to ascertain an approximation of their emotional state. Using electrical muscle stimulation their face is animated into an expression that corresponds with the output of the emotional recognition system. Through our work, we contribute the insight of three possible dark patterns that might emerge from emotional integration, including: reductionism of human emotion, disruptions of agency, and parasitic symbiosis. We hope that these insights inspire researchers and practitioners to approach human-computer integration more cautiously.

CCS CONCEPTS

• **Media Arts**; • **Human-computer Interaction (HCI)**; • **Interaction Design**;

KEYWORDS

Media Art, Brain-computer interface, EEG, Electrical Muscle Stimulation, Emotion detection

ACM Reference Format:
Rod Dickinson, Nathan Semertzidis, and Florian 'Floyd' Mueller. 2022. Machine In The Middle: Exploring Dark Patterns of Emotional Human-Computer Integration Through Media Art. In *CHI Conference on Human Factors in Computing Systems Extended Abstracts (CHI '22 Extended Abstracts), April 29-May 5, 2022, New Orleans, LA, USA.* ACM, New York, NY, USA, 7 pages. https://doi.org/10.1145/3491101.3503555

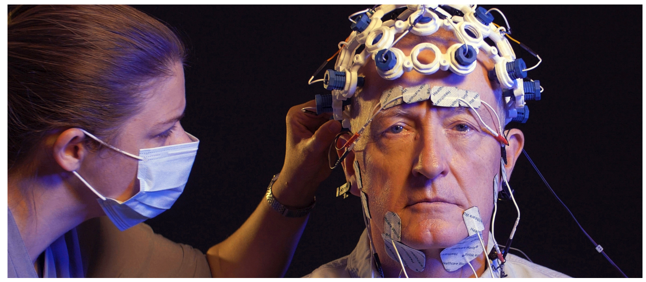

Figure 1: machine_in_the_middle, a media art piece that uses brain-computer interfacing and electrical muscle stimulation to explore the dark patterns of emotional human-computer integration.

1 INTRODUCTION

There has been a long history of attempts to quantify, categorize, operationalise and classify human emotion. Examples of this stretch as far back as Darwin's work "Expression of the Emotions in Man and Animals" [15]; Ekman's highly contested hypothesis that emotions have a universal form and expression [19]; the scientific lens through which psychology and neuroscience operationalise emotion[10]; and perhaps most recently, the contemporary growing area of research within HCI known as "affective computing" [44]. In each of these cases, there is an attempt to render emotions visible and numerically translatable such that they can be ultimately integrated with machinic processes.

Underlying the majority of these endeavours is the assumption that the translation of the subjective experience of emotion to an objective classification is possible and desirable. Advocates argue these processes will not only help understand ourselves better, but also be instrumental in progress toward a more humanist future where machines may be able to interpret human emotion and respond accordingly [11, 52]. However, for each antecedent step forward in classifying human emotion, we see a consequent "dark pattern" bifurcate, with dark patterns referring to instances where user value is supplanted for another interest group [24, 25]. For

CHI '22 Extended Abstracts, April 29-May 5, 2022, New Orleans, LA, USA

ACM ISBN 978-1-4503-9156-6/22/04.
https://doi.org/10.1145/3491101.3503555

instance, Darwin's work in cataloging emotional expression went on to serve as the foundation for the misguided and sexist concept of "hysteria" in early psychology [11, 51]. Similarly, Ekman's work on the universal classification of emotions that underlie human facial expressions was appropriated by US airport security to identify potential terrorists, resulting in racial profiling, ultimately strengthening racism in the US national security state [54]. Advances in affective computing have also been criticized with accusations that the entanglement of psychology and computer science has led to the "calculability of human subjectivity", quantifying humans into information for psychographic models through which individuals can be digitally categorised [49, 50].

With contemporary advances in affective technology the dialectic between well-intentioned aims and unintended nefarious consequences become ever more critical: Specifically, as the emerging HCI paradigm of "Human-Computer Integration" purports, humans and computers are increasingly moving toward a trajectory in which the two will inevitably merge into a symbiotic organism [20–22, 36, 38, 39]. In this respect, affective computing can be seen as a shift away from the interpretation and categorization of facial expressions, behaviours, and observable affective cues, and toward the incorporation of users' neurophysiology and potentially the seat of emotion itself [6, 46]. Through technologies such as electroencephalography and electric muscle stimulation, the machine is now able to bidirectionally sense and modulate neural activity as if it was itself part of the user's body [35, 37]. Proponents of this perspective point to the future of HCI as unequivocally positing human-computer integration as "inevitable, necessary, and desirable [21]". Yet in considering the "dark patterns" that have historically emerged from classifying emotions, some serious concerns are raised: what if in intending to design an emotional symbiosis with machines that is mutualistic (where both the human and machine benefit), we instead create a parasitic symbiosis, where the user is subsumed for the machine?

In this work, we explore the emergence of this dark pattern by preemptively actualising it through interactive technology, as others in HCI have done before us when studying dark patterns in other HCI contexts [17, 24, 25]. We do so through our work "machine_in_the_middle", an interactive system that viscerally realises the potential consequences of asking a machine to direct our behaviour based on its interpretation of what we are feeling. Machine_in_the_middle bidirectionally actuates the user's neurophysiological activity using electroencephalography (EEG) to capture the subject's neural activity and uses electrical muscle stimulation to control the expression of the subject's face. We present this exploration as an artwork, first due to art's ability to provoke and speculate future technological and cultural trajectories, but also in reference to the historical association art has had with emotion and human emotional expression. Through our work, we contribute prescient insights into possible dark patterns that might emerge through emotional symbiosis with machines, and hope that these insights inspire future researchers and practitioners to approach human-computer integration more cautiously.

2 RELATED WORK

Machine_in_the_middle is both inspired by, and contextualised within, a rich cannon of experimental artwork and HCI research. Here we discuss the related work which has informed it.

2.1 Art

Machine_in_the_middle sits between the fields of conceptual art and affective computing. It uses EEG data as a human computer interface and as such has precursors in many artworks that also utilise EEG processing as brain-computer interfaces. These include relatively recent artworks such as Suzanne Dikker's collaboration with Marina Abramovic "Measuring the Magic of Mutual Gaze" [32] which explores the synchronicity of EEG signals between two subjects and "E.E.G. Kiss" by Karen Lancel and Hermen Maat [31] which investigates how a kiss can be translated into bio-feedback data and visualised. Older artworks include Alvin Lucier's seminal "Music for Solo Performer" [7] where a performer triggers musical instruments using alpha brain waves to create an experimental score. In each of these examples, and many more, the materialisation of EEG data (visually, audibly or both) is the tangible, artistic output and is often displayed in dramatic form, such as "E.E.G. Kiss" where the EEG waveform is projected around the performers. In these examples EEG data flows from human participant to the machine and is then materialised and as such almost analogous to other more traditional expressive art forms, such as painting. Machine_in_the_middle breaks this methodology by creating a feedback loop with the machine and subject, where the latter involuntarily materialises the machine's interpretation of the EEG data on their face.

Despite claims that the techniques deployed by these cited examples have been adopted from Cybernetics [7, 47], the feedback loop created by machine_in_the_middle is far closer to an approximation of a first order Cybernetic system, characterised by a closed feedback loop, and much less reliant on the aesthetic and expressive potential of the captured EEG data, which despite the deployment of cutting edge technology, is rooted in a idea of "expressive art-making" that derives from the romantic era [45].

By contrast, other artworks that utilise the relationship between body and machine as a literal and conceptual feedback loop, much in the same way as machine_in_the_middle, include Pedro Lopes' "Ad Infinitum" [34] where a "parasitical" machine "attaches electrodes onto the human visitors harvesting their kinetic energy by electrically persuading them to move their muscles using EMS" (Electrical Muscle Stimulation) [37] and Sarah Selby's 2019 installation "Raised by Google" [1] where visitors to the installation are psychometrically profiled with software designed by Cambridge University's Psychometric Centre and then directed through a giant, physical maze of rooms, such that the software is in control of the visitor's movements.

In both these artworks, and in machine_in_the_middle, the classical HCI configuration, in which the human is always in control, is reversed. Instead, in these artworks the participants experience how it feels when a machine is in control.

2.2 HCI Systems

As machine_in_the_middle itself is a functioning prototypictical system, the work has been inspired and informed by previous HCI research, in particular works which contain human-computer integration or involve the sensing and stimulation of human neurophysiology.

One of such works is Neo-Noumena [46]. Neo-Noumena is a wearable system that uses EEG and mixed reality to augment interpersonal emotion communication. While wearing the system, users have their emotions algorithmically classified via EEG data in real-time. Based on the system's interpretation of their emotional state, a procedurally generated flock of fractals representing their emotional state is generated in augmented reality (AR) that can be seen by other users. Evaluation of the system's user experience suggested participants found Neo-Noumena to be a useful tool for discussing and exploring the emotional experiences of their partner, while also allowing for the interpersonal regulation of emotion. Another similar work is "Wigglears" [42], a wearable system that wiggles the user's ears based on their emotion. Specifically, the system interprets the skin conductance of its user to classify the user's emotional state. Based on the classification, the user's ears are mechanically "wiggled" by servo motors mounted behind the user's ears, creating an emotionally expressive gesture similar to the emotionally communicative ear movements of animals such as dogs and cats.

Taken together, machine_in_the_middle builds off these works in that it similarly classifies human emotion and then uses this information to produce associated emotionally representative output. However, while these previous systems adapt technologies or body parts as design materials otherwise not connected to emotion, machine_in_the_middle differs in that it employs the human face, a space already used by us to communicate emotion, to express the machines interpreted classified emotional states. Through this approach, we hope to highlight the juxtaposition between the human experience of emotion, and the machine's interpretation of this experience and subsequent attempt at an authentic replication. As such, we also draw inspiration from work in HCI that have used electrical muscle stimulation (EMS) to animate the body.

One of such works employing EMS is "muscle plotter" [37]. Muscle plotter is a wearable system that delegates control of the user's arm to a computer to allow the computer to use the arm as a printing mechanism. In their work, the authors give an example scenario of an engineering car sketch. The user sketches their car and then attempts to illustrate the car's aerodynamics by drawing airflow lines around the car. While such illustrations are typically too complex for the human to complete freehand, requiring complex mathematics, the user is able to delegate control of their arm to the system, which completes the drawing vicariously through them via EMS. In another study using EMS titled "Eyewear to Make Me Smile" [56], the authors develop a wearable system that uses EMS to stimulate the wearer's facial expression into a smile. The system was inspired by the James-Lang theory of emotion, which posits that physiological actions such as smiling can alter emotional experiences [12].

In considering these works together, we acknowledge EMS as an effective means to animate the body, create experiences of giving up control of the body to the machine, and to also generate emotional expressive facial expressions. With these considerations in mind, we employed EMS in the design of machine_in_the_middle.

3 MACHINE_IN_THE_MIDDLE

Machine_in_the_middle was made for an online exhibition curated by Cognitive Sensations, an arts organisation based in Liverpool, UK.

The artwork is a 4-minute demo and explanatory video of a novel BCI system that creates a feedback loop between subject and machine, explicitly giving the machine control over the subject's body. Thematically the artwork was part of a series of events that explored the interrelationship between computing and the brain.

Figure 2: The setup used for machine_in_the_middle. On the left Steve Davis is seated, equipped with the EEG headset and adorned with EMS electrodes on his face. To the right is the computer system processing and visualising Steve's brain activity while interpreting his emotions and controlling the EMS device.

3.1 Designing Machine_in_the_Middle

Machine_in_the_middle is a reimagining of a "Man (or machine) in the Middle" cyber-attack. It was designed as a functioning prototype: A hardware and software system that captures and determines the emotional state of a subject and gives the system control over the subject's facial expressions.

Typically a Man in the Middle attack involves a hidden perpetrator who secretly intercepts the communications between two parties, either to eavesdrop, or to modify traffic travelling between them. In this case the two parties are the brain and the face of a single subject and the Man (or machine) in the Middle is the hardware and software system that hijacks the communication between the subject's brain and face.

In this type of cyber-attack the victims lose control of their ability to communicate unhindered. Man in the Middle attacks are implemented through the interception and the decryption of communication which denies the victim agency and control. Our artwork, machine_in_the_middle, uses this as a deliberately provocative paradigm to explore our increasingly intimate relationship with technology. It inverts the normal understanding of HCI where the human subject is in control, instead placing the machine in control. The machine decides how our subject facially expresses themselves. Machine_in_the_middle renders visible a new kind of relationship with technology where machines play a powerful role, creating new human-machine assemblages and new subjectivities, in what Amoore calls a new kind of "we", with "new limits and thresholds of what it means to be human" [5].

Figure 3: The EMS stimulation contorts Steve Davis's mouth into a not-so-authentic expression of stress

Our video demoing the machine_in_the_middle system shows it being tested on World Snooker Champion Steve Davis. Davis was the perfect subject because during his period of multiple World Championship wins in the 1980s and 1990s he was famous for his deadpan demeanor, never openly expressing any emotion during snooker tournaments [26]. The design of machine_in_the_middle overrode any tendency Davis might have had to not facially express his emotions.

3.2 Technical Details of Machine_in_the_Middle

The technical execution of machine_in_the_middle can be reduced to two major components, the BCI-driven emotion classification, and Arduino-controlled electrical muscle stimulation.

3.2.1 Emotion Classification. The system interprets EEG data to classify the participants' concurrent emotional state in real-time. To achieve this, the brain activity of the participant is sensed through an eight channel OpenBCI Ultracortex Mark IV 3D printed headset, with the electrophysiological signal amplified by an OpenBCI Cyton board [4]. The electrode placement of the headset employed the 10-20 EEG electrode placement convention [28], with the chosen recording sites being: Fp1, Fp2, F3, F4, F7, F8, T3, T4, with A1 and A2 used as ground and reference respectively. The choice of electrode sites was informed by the electrode placement utilised in the LUMED data set collection study [13, 14], the dataset we used for training the emotion classifier (described in detail below). From the Cyton board, EEG data is streamed via bluetooth to a nearby laptop at a sample rate of 250Hz.

Once received by the laptop, the EEG data is processed and classified via a custom python script using the "brainflow" python API. Data is first filtered through a 0.5 - 59 Hz Bessel bandpass filter [53], with an additional 50Hz butterworth filter for removing ambient electronic noise [3]. To further mitigate environmental noise and myoelectric artefact, a "coif3" wavelet denoising filter was employed [29]. Once the signal is filtered and denoised, the power spectral density (PSD) is calculated for each channel using the Welch method [48] with a Blackman-Harris windowing function that has an overlap of three samples [30]. The band powers for bandwidths 1-4hz (delta), 5-8hz (theta), 9-13hz (alpha), 14-30hz (beta), and 31-50hz (gamma) are then calculated from the PSD for each channel individually.

To obtain a classification of valence, band power values are binned in five second moving windows and are then fitted to a support vector machine emotion classifier which classifies the data as either "high valence" (positive emotion) or "low valence" (negative emotion). For building the classifier, we used the "scikit-learn" support vector machine "support vector classifier" with "balanced" class weights [41], which produced an accuracy of 63.5%, trained with the LUMED emotion data set. To obtain a classification of arousal, band power values are binned in five second moving windows and then fitted to brainflows pre-built relaxation regression model. The "relaxation" metric is then normalised and inverted, transforming it to an "arousal" metric. Arousal values greater than 0.5 return a classification of "high arousal", and arousal values of less than 0.5 return a classification of "low arousal". Classifications of valence and arousal are then considered together to provide a two-dimensional classification of emotion. Each pairing of valence and arousal classification produces a unique emotional category, which ultimately add up to a set of four possible emotions the system can classify brain activity under. The dimensional parings and subsequent categories are illustrated below in figure 3.

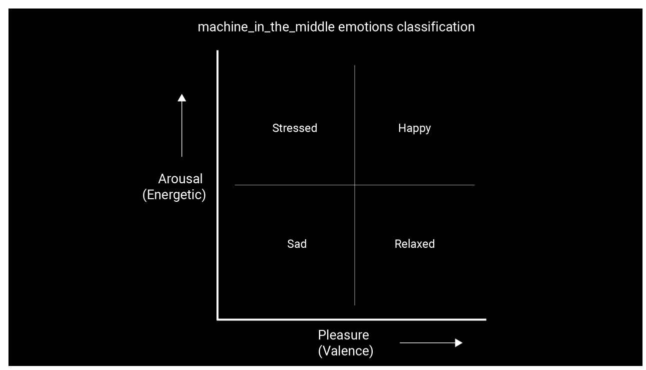

Figure 4: The 2D Valence-Arousal model of emotion and the system's possible emotion classification classes.

3.2.2 Arduino-Controlled Electrical Muscle Stimulation. To animate the facial expression of the participant, machine_in_the_middle uses 3 transcutaneous electrical nerve stimulation (TENS) machines to deliver electrical muscle stimulation to the face via 12 electroconductive sticky pads. The 12 pads vary in size and are strategically placed around the participant's jaw, chin, and forehead, such that the system has access to a wide enough range of facial muscles to produce stressed, happy, sad, and relaxed facial expressions. The placement of the pads and the voltage setting was critical to target each muscle accurately. While the participant wears the system, the TENS machines are on but stimulation is cut at a relay board which all the pads are connected to. Once the emotion classifier classifies an emotion, the classification is sent to an Arduino controlling the relay board, which then opens the circuit between the TENS machine and the pads associated with the desired facial expression, contracting the muscles and animating the face.

4 REFLECTION

We are now in a position to offer a number of insights arising from our reflection on machine_in_the_middle. Specifically, these insights are drawn from our knowledge of designing, crafting and developing the system, our observations from deploying the system and having Steve Davis use it, the reactions of those who viewed the artwork, and the formal discussions during the launch event of the artwork at the "Cognitive Sensations" virtual exhibition and event "Decoding Humans". Our insights are framed as three overarching conceptual themes: Reduction; Agency; and Symbiosis. In presenting these themes, we also discuss their connection to broader related theory.

4.1 Reduction: Emotion As Number

The ever increasing intimacy of computing and computers, from mobile to wearables, has made the possibility of computers sensing users' emotional states ever closer and ever less reliant on users' externalising their moods via mood tracking apps and GUIs. Machines increasingly and directly retrieve data about the affective state of their user in real time, even informing them how they felt [16] or how 'well' they are. Typically the data is reduced to a set of numerical representations.

For machine_in_the_middle to function it was necessary to subject the many factors that constitute the experience of emotion to a similar gross computational reduction. This initially takes place when the hardware and software determine the emotional state of the subject with the experience of emotion being reduced to voltage differentials, which are then in turn further reduced through a series of filters and data transformations, which ultimately limit the variation of emotional experience to four categories. Thus, it might be more appropriate to conceptualise this numerical transformation as a new type of emotion, an "emotion-as-number" [18] and to consider it as parallel to the emotional states we experience at a human visceral level. Much in the way that we understand a social media 'like' or 'friend' as a parallel to actually liking something or knowing someone.

Further numerical reduction was designed into machine_in_the_middle to articulate the emotional expression of our subject's face: 12 of the face's 42 muscles were chosen to express only four possible emotions through binary "on-or-off" electrical stimulation. At each step, the richness of the experience of emotion is sacrificed for the machine to function, reducing the human to operate on the machines' terms, as Wiener proposed in his early formulation of the theory of cybernetics [23]. Machine_in_the_middle intentionally explores these parallel, reduced, numerical emotions as a way of preempting the emergence of "emotion-as-number". It does this by producing a visceral demo, showing how an affective experience can be rendered emotion free; or alternatively, how it can be abstracted and attenuated, such that emotions can now be included in calculations about efficiency and effectiveness [18]. This totalising approach promises a computational rendering of all human behaviour [43].

Categorical biosensing systems may bias and reduce our understanding of emotional experiences to what is quantifiable and machine interpretable, with nuance being conditioned out of our spectrum of emotional experience. Similar observations have been made demonstrating that users often consult biodata-driven classifications of emotions as objective arbiters of emotional truth, believing that if a conflict between the machines' classification and their own experience emotion arises from user error, this may lead them to question "is there something wrong with me?" [8, 9, 46].

While these examples demonstrate emotional reductionism when engaging with digital media, machine_in_the_middle hints at a future in which the user's ontology and entire emotional reality may be radically shifted due to the omnipresent nature of the technology and its intimate connection with their body. Will emotion integration systems ultimately narrow the complexity, depth and nuance of human emotions to a degree that some emotions will ultimately become invalidated or unrecognizable within a received set of technologically driven systems?

4.2 Agency: Facial Expression And Intentionality

The real time monitoring of physiological data, alongside facial recognition systems, might appear to promise an empathetic media [2, 40, 55] that connects the user's physiological data with their outward expression. Both Tomkins and Ekman subscribed to this idea positing that facial expressions were an involuntary expression of the inner state of a subject, a hard-wired response to a stimulus that triggers it [33] rather than an intentional response. Thus, they concluded it was reasonable to make assumptions about a subject's state of mind from their facial expression. Their critics contend that humans produce facial expressions that are often intentionally deceptive, particularly when it is advantageous for them and that Tomkins and Ekland's model of behavioural responses seeks to make a simplistic and false equation between facial expression and emotional states. Nevertheless, Ekman's model has been widely adopted and used, even by his critics [33].

Equally controversially Tomkins and Ekland believed that there were pancultural emotional expressions and Tomkins developed an observational system "The Facial Action Scoring Technique" to categorise them. This approach was also indebted to a rationalist, cybernetic view of human behaviour, in which subjects constantly adjust their behaviour as part of a real time feedback loop [16, 27]. Machine_in_the_middle formulated a similar feedback loop between the brain and face of our subject, Steve Davis, but framing it as an attack on him rather than the creation of a rational or empathetic system. The attack proceeds by our machine_in_the_middle system removing Davis' ability to control his facial expressions, instead our system took control of his facial muscles through electrical muscle stimulation. This has a two-fold effect: First, the amount of emotional information able to be conveyed in his expression is reduced to a set of basic expressions or categories of stressed, happy, sad, and relaxed. Second, the ability for Davis to control his facial expressions in order to conceal or intentionally misinform others of his emotional state is removed.

The machine_in_the_middle system mirrors the dark pattern in the real-time motoring of emotional physiology ultimately disrupting the privacy or even agency of the individual under the guise of increasing empathy and transparency.

4.3 Symbiosis: Mutualistic Integration Versus Parasitic subsumption

The theoretical cannon of human-computer integration often describes the ongoing merger with the human body and computational machines as a form of emerging symbiosis. However, much of this work only frames that symbiosis as mutualistic and omits the fact that in nature there are several types of symbiotic relationships: As well as mutualistic symbiosis, where both organisms benefit, symbiotic relationships between organisms can be further described as: commensalistic, in which one organism benefits while the other is unaffected; and parasitic, where in which one organism benefits while the other suffers. This raises the question, what kind of symbiosis does machine_in_the_middle represent? In machine_in_the_middle, the machine is the obvious benefactor, with the human sacrificing much to ensure its operation. This sacrifice comes in the form of a radically devolved experience of emotion as well as a disruption to the human agency, as described earlier. However adding to this is the observation that there is a notable expression of pain in the face of Steve Davis that subtly underlies his artificially articulated expression. While this may appear to be a superficial observation when considering that machine_in_the_middle was not designed with utilitarian goals in mind, it speaks to other EMS devices and their associated studies that do present themselves as tools for good. Perhaps the strongest example of this would be "Eyewear to Make Me Smile" [56], which we discussed earlier. The authors present their work in which they use EMS to forcefully articulate a smile on the face of the participant to make them happier. However, through developing machine_in_the_middle we have come to understand facial EMS to be an unpleasant experience.

Through these observations we identify a dark pattern in the design of symbiotic systems that may be best described as parasitic, despite intentions to create the opposite. Often this is paired with an oversight of the cost to human experience. Ironically the very systems intended to humanise technologies have resulted in alienating experiences. Others have also similarly described how categorised expressions of emotion (such as social media emoji reactions) do not benefit humans but rather the social media platforms that they are deployed on [50]. This ultimately creates a form of symbiosis in which human emotion is objectified in order to sustain the cybernetic system of surveillance capitalism as the subject is objectified, digitised, tracked, and guided through algorithmically steered "user journeys" [50].

We argue that human-computer integration does not need to replicate these mistakes in the future. The rationale for this argument is built on two important factors. On the one hand, the parasitic tendencies of symbiotic systems stem, arguably, from their foundations in Wiener's conceptualisation of cybernetics [23]. In Wiener's understanding, cyber-physical systems oblige the human subject to be reductively codified in order to be machine interpretable. Cybernetics, being fundamentally mathematical, needs, in other words, to build 'up' to the level of phenomena. On the other hand however, and in contrast with Wiener's conceptualisation, human-computer integration is fundamentally phenomenological: though it also deals with human-machine systems, it begins with human experience, before moving 'down', where necessary, to mathematical mechanics. The argument here is that, if cybernetics functions by bringing humans 'down' to the level of the machine, human-computer integration should strive to bring the machine 'up' to the level of the human.

5 CONCLUSION

In this work, we propose that there is historically a dark pattern which emerges when humans have consistently attempted to classify emotion with good intentions, only for nefarious unintended consequences to emerge from these efforts. With this in mind, we draw attention to the growing symbiosis between the computational machine and the human body, as described by the HCI paradigm of human-computer integration, and argue that it is especially critical that we anticipate the dark patterns that may arise from such technologies given their very intimate relation with the human body, directly answering the call for an exploration of dark-paths in human-computer integration made by Mueller et al. [38]. As such, we explore this dark pattern through our work machine_in_the_middle, an interactive system that viscerally realises the potential consequences of asking a machine to direct our behaviour based on its interpretation of what we are feeling. Machine_in_the_middle bidirectionally actuates the user's neurophysiological activity using EEG to capture the subject's neural activity and uses electrical muscle stimulation to control the expression of the subject's face, creating an immediate triangular feedback loop between the subjects EEG data, the computational system and electrical muscle stimulation that controls their facial muscles. We then reflect on this work and contribute prescient insights into possible dark patterns that might emerge through emotional symbiosis with machines, and hope that these insights inspire future researchers and practitioners to approach human-computer integration more cautiously.

ACKNOWLEDGMENTS

Machine_in_the_middle was created by Rod Dickinson and Nathan Semertzidis for The Downloadable Brain season of events produced by Cognitive Sensations (https://www.cognitivesensations.com/). All images are by Helm Films, Paul Blakemore, Rod Dickinson and Nathan Semertzidis.

REFERENCES

[1] [n.d.]. Raised by Google. https://www.arebyte.com/raised-by-google
[2] Andrea F Abate, Aniello Castiglione, Michele Nappi, and Ignazio Passero. 2020. DELEX: a DEep Learning Emotive eXperience: Investigating empathic HCI. In *Proceedings of the International Conference on Advanced Visual Interfaces*. 1–8.
[3] Imteyaz Ahmad, Faruk Ansari, and UK Dey. 2012. A review of EEG recording techniques. *International Journal of Electronics and Communication Engineering & Technology (IJECET)* 3, 3 (2012), 177–186.
[4] Audrey Aldridge, Eli Barnes, Cindy L Bethel, Daniel W Carruth, Marianna Kocturova, Matus Pleva, and Jozef Juhar. 2019. Accessible electroencephalograms (EEGs): A comparative review with openbci's ultracortex mark IV headset. In *2019 29th International Conference Radioelektronika (RADIOELEKTRONIKA)*. IEEE, 1–6.
[5] Louise Amoore. 2020. *Cloud ethics*. Duke University Press.
[6] Josh Andres, mc schraefel, Nathan Semertzidis, Brahmi Dwivedi, Yutika C Kulwe, Juerg von Kaenel, and Florian Mueller. 2020. Introducing Peripheral Awareness as a Neurological State for Human-computer Integration. In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems*. 1–13.
[7] G Douglas Barrett. 2017. The Brain at Work: Cognitive Labor and the Posthuman Brain in Alvin Lucier's Music for Solo Performer. *Postmodern Culture* 27, 3 (2017).

- [8] Kirsten Boehner, Rogério DePaula, Paul Dourish, and Phoebe Sengers. 2005. Affect: from information to interaction. In *Proceedings of the 4th decennial conference on Critical computing: between sense and sensibility*. 59–68.
- [9] Kirsten Boehner, Rogério DePaula, Paul Dourish, and Phoebe Sengers. 2007. How emotion is made and measured. *International Journal of Human-Computer Studies* 65, 4 (2007), 275–291.
- [10] Joan C Borod et al. 2000. *The neuropsychology of emotion*. Oxford University Press.
- [11] Rafael A Calvo, Sidney D'Mello, Jonathan Matthew Gratch, and Arvid Kappas. 2015. *The Oxford handbook of affective computing*. Oxford Library of Psychology.
- [12] Walter B Cannon. 1927. The James-Lange theory of emotions: A critical examination and an alternative theory. *The American journal of psychology* 39, 1/4 (1927), 106–124.
- [13] Yucel Cimtay and Erhan Ekmekcioglu. 2020. Investigating the use of pretrained convolutional neural network on cross-subject and cross-dataset EEG emotion recognition. *Sensors* 20, 7 (2020), 2034.
- [14] Yucel Cimtay, Erhan Ekmekcioglu, and Seyma Caglar-Ozhan. 2020. Cross-subject multimodal emotion recognition based on hybrid fusion. *IEEE Access* 8 (2020), 168865–168878.
- [15] Charles Darwin. 2015. *The expression of the emotions in man and animals*. University of Chicago press.
- [16] William Davies. 2017. How are we now? Real-time mood-monitoring as valuation. *Journal of Cultural Economy* 10, 1 (2017), 34–48.
- [17] Linda Di Geronimo, Larissa Braz, Enrico Fregnan, Fabio Palomba, and Alberto Bacchelli. 2020. UI dark patterns and where to find them: a study on mobile applications and user perception. In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems*. 1–14.
- [18] Otniel E Dror. 2001. Counting the affects: Discoursing in numbers. *Social research* (2001), 357–378.
- [19] Paul Ekman and Dacher Keltner. 1997. Universal facial expressions of emotion. *Segerstrale U, P. Molnar P, eds. Nonverbal communication: Where nature meets culture* 27 (1997), 46.
- [20] Umer Farooq and Jonathan Grudin. 2016. Human-computer integration. *interactions* 23, 6 (2016), 26–32.
- [21] Umer Farooq, Jonathan Grudin, Ben Shneiderman, Pattie Maes, and Xiangshi Ren. 2017. Human computer integration versus powerful tools. In *Proceedings of the 2017 CHI Conference Extended Abstracts on Human Factors in Computing Systems*. 1277–1282.
- [22] Umer Farooq and Jonathan T Grudin. 2017. Paradigm shift from human computer interaction to integration. In *Proceedings of the 2017 CHI Conference Extended Abstracts on Human Factors in Computing Systems*. 1360–1363.
- [23] Peter Galison. 1994. The ontology of the enemy: Norbert Wiener and the cybernetic vision. *Critical inquiry* 21, 1 (1994), 228–266.
- [24] Colin M Gray, Yubo Kou, Bryan Battles, Joseph Hoggatt, and Austin L Toombs. 2018. The dark (patterns) side of UX design. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems*. 1–14.
- [25] Saul Greenberg, Sebastian Boring, Jo Vermeulen, and Jakub Dostal. 2014. Dark patterns in proxemic interactions: a critical perspective. In *Proceedings of the 2014 conference on Designing interactive systems*. 523–532.
- [26] Phil Haigh. 2020. Ronnie O'Sullivan: Steve Davis ruined snooker in the '80s. https://metro.co.uk/2020/12/09/ronnie-osullivan-steve-davis-ruined-snooker-in-the-80s-13724413/
- [27] Orit Halpern. 2014. Cybernetic rationality. *Distinktion: Scandinavian Journal of Social Theory* 15, 2 (2014), 223–238.
- [28] Richard W Homan, John Herman, and Phillip Purdy. 1987. Cerebral location of international 10–20 system electrode placement. *Electroencephalography and clinical neurophysiology* 66, 4 (1987), 376–382.
- [29] Saleha Khatun, Ruhi Mahajan, and Bashir I Morshed. 2016. Comparative study of wavelet-based unsupervised ocular artifact removal techniques for single-channel EEG data. *IEEE journal of translational engineering in health and medicine* 4 (2016), 1–8.
- [30] Andi Kivinukk and Gert Tamberg. 2007. On Blackman-Harris Windows for Shannon Sampling Series. *Sampling Theory in Signal & Image Processing* 6, 1 (2007).
- [31] lancel/maat. [n.d.]. E.E.G. KISS. https://www.lancelmaat.nl/work/e.e.g-kiss/
- [32] Ellen K Levy, Robert Beck, Jennifer Bornstein, Suzanne Dikker, Matthias Oostrik, Greg Garvey, Nicole Ottiger, Jane Philbrick, and Jill Scott. 2014. LEONARDO GALLERY: Sleuthing the Mind. *Leonardo* (2014), 427–458.
- [33] Ruth Leys. 2010. How did fear become a scientific object and what kind of object is it? *Representations* 110 (2010), 66–104.
- [34] Pedro Lopes. 2018. ad infinitum. https://a-parasite.org/
- [35] Pedro Lopes. 2018. The next generation of interactive devices Human Computer Interaction Lab, Hasso Plattner Institute. *XRDS: Crossroads, The ACM Magazine for Students* 24, 3 (2018), 62–63.
- [36] Pedro Lopes, Josh Andres, Richard Byrne, Nathan Semertzidis, Zhuying Li, Jarrod Knibbe, Stefan Greuter, et al. 2021. Towards understanding the design of bodily integration. *International Journal of Human-Computer Studies* 152 (2021), 102643.
- [37] Pedro Lopes, Doăa Yüksel, François Guimbretière, and Patrick Baudisch. 2016. Muscle-plotter: An interactive system based on electrical muscle stimulation that produces spatial output. In *Proceedings of the 29th Annual Symposium on User Interface Software and Technology*. 207–217.
- [38] Florian Mueller, Pedro Lopes, Paul Strohmeier, Wendy Ju, Caitlyn Seim, Martin Weigel, Suranga Nanayakkara, Marianna Obrist, Zhuying Li, Joseph Delfa, et al. 2020. Next Steps for Human-Computer Integration. In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems*. 1–15.
- [39] Florian Mueller, Pattie Maes, and Jonathan Grudin. 2019. Human-Computer Integration. (2019).
- [40] Samuel Marcos Pablos, Jaime Gómez García-Bermejo, Eduardo Zalama Casanova, and Joaquín López. 2015. Dynamic facial emotion recognition oriented to HCI applications. *Interacting with Computers* 27, 2 (2015), 99–119.
- [41] Fabian Pedregosa, Gaël Varoquaux, Alexandre Gramfort, Vincent Michel, Bertrand Thirion, Olivier Grisel, Mathieu Blondel, Peter Prettenhofer, Ron Weiss, Vincent Dubourg, et al. 2011. Scikit-learn: Machine learning in Python. *the Journal of machine Learning research* 12 (2011), 2825–2830.
- [42] Victoria Peng. 2021. Wigglears: Wiggle Your Ears With Your Emotions. In *Extended Abstracts of the 2021 CHI Conference on Human Factors in Computing Systems*. 1–5.
- [43] Alex Pentland. 2014. *Social physics: How good ideas spread-the lessons from a new science*. Penguin.
- [44] Rosalind W Picard. 1999. Affective Computing for HCI.. In *HCI (1)*. Citeseer, 829–833.
- [45] Herbert Read and Sreten Marić. 1949. *The meaning of art*. Vol. 213. Penguin books Bungay, Suffolk, England.
- [46] Nathan Semertzidis, Michaela Scary, Josh Andres, Brahmi Dwivedi, Yutika Chandrashekhar Kulwe, Fabio Zambetta, and Florian Mueller. 2020. Neo-Noumena: Augmenting Emotion Communication. In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems*. 1–13.
- [47] Mohammad Shidujaman. 2020. Mapping Between Mind Cybernetics and Aesthetic Structure in Real-Time EEG Art. In *HCI International 2020-Late Breaking Papers: Multimodality and Intelligence: 22nd HCI International Conference, HCII 2020, Copenhagen, Denmark, July 19-24, 2020, Proceedings*, Vol. 12424. Springer Nature, 344.
- [48] OM Solomon Jr. 1991. PSD computations using Welch's method. *NASA STI/Recon Technical Report N* 92 (1991), 23584.
- [49] Luke Stark. 2018. Algorithmic psychometrics and the scalable subject. *Social Studies of Science* 48, 2 (2018), 204–231.
- [50] Luke Stark and Kate Crawford. 2015. The conservatism of emoji: Work, affect, and communication. *Social Media+ Society* 1, 2 (2015), 2056305115604853.
- [51] Frank J Sulloway. 1992. *Freud, biologist of the mind: Beyond the psychoanalytic legend*. Harvard University Press.
- [52] Jianhua Tao and Tieniu Tan. 2005. Affective computing: A review. In *International Conference on Affective computing and intelligent interaction*. Springer, 981–995.
- [53] Lorenz Trachsel. 1993. Hartley transforms and narrow bessel bandpass filters produce similar power spectra of multiple frequency oscillators and all-night EEG. *Sleep* 16, 6 (1993), 586–594.
- [54] Sharon Weinberger. 2010. Airport security: Intent to deceive? *Nature News* 465, 7297 (2010), 412–415.
- [55] Peter Wright and John McCarthy. 2008. Empathy and experience in HCI. In *Proceedings of the SIGCHI conference on human factors in computing systems*. 637–646.
- [56] Lai Yen-Chin, YuanLing Feng, Kai Kunze, Junich Shimizu, and Takuro Nakao. 2017. Eyewear to Make Me Smile: Can Electric Muscle Stimulation increase Happiness?. In *Proceedings of the Eleventh International Conference on Tangible, Embedded, and Embodied Interaction*. 579–582.