Facing Forward: Policy for Automated Facial Expression Analysis

Jeffrey K. Mullins
University of Arkansas, jmullins@walton.uark.edu

Patrick A. Stewart
University of Arkansas, pastewar@uark.edu

Thomas J. Greitens
Central Michigan University, thomas.greitens@cmich.edu

Facing Forward: Policy for Automated Facial Expression Analysis

Jeffrey K. Mullins,[1] Patrick A. Stewart,[2] Thomas J. Greitens[3]
[1]University of Arkansas, USA, jmullins@walton.uark.edu
[2]University of Arkansas, USA, pastewar@uark.edu
[3]Central Michigan University, USA, thomas.greitens@cmich.edu

Abstract

The human face is a powerful tool for nonverbal communication. Technological advances have enabled widespread and low-cost deployment of video capture and facial recognition systems, opening the door for automated facial expression analysis (AFEA). This paper summarizes current challenges to the reliability of AFEA systems and challenges that could arise as a result of reliable AFEA systems. The potential benefits of AFEA are considerable, but developers, prospective users, and policy makers should proceed with caution.

Keywords: Facial Expression Analysis, Facial Recognition, Emotion, Policy

1 Introduction

The face conveys identity, which is why facial photographs have long been used on driver's licenses and passports. The face can also reveal internal states of mind and behavioral intentions. Identity-oriented facial recognition is increasingly used in security, social media image tagging, and personal device access. With advances in computing, software, high-definition video, and measuring human physiology, machines are beginning to recognize what people feel and think. If *automated facial expression analysis* (AFEA) were to reliably reveal emotional and cognitive states, it could be used to improve organizational performance, individual well-being, and public safety. For example, AFEA-enabled machines could be used to identify genuinely excited job candidates, frustrated customers in a transaction, or intensely angry travelers going through security checkpoints. However, problems arise if the technology does not work and also if it does.

This policy editorial "looks down the road." It starts with identity analysis, which is happening now and is

itself controversial, and then moves on to assess roadblocks and other hazards regarding AFEA. We offer guidelines for creators of AFEA systems, as well as for those hoping to implement AFEA as part of their organizational processes. Potential policy issues in practice are summarized in Table 1 and discussed below as *challenges to reliability* (i.e., they could cause the technology not to work) and *challenges of reliability* (i.e., problems that arise if the technology *does* work).

2 AFEA Overview

The concept of identity is fairly straightforward, and photos have long been part of identity. Computer-assisted facial recognition is an extension of this but is controversial if it is involved in imposing sanctions (e.g., arrest) or granting rewards. Generally, facial recognition for identity purposes that do not involve sanctions or rewards is not subject to policy scrutiny. In cases involving sanctions or rewards, the issues can generally be sorted out using relatively well-established risk analysis.

Table 1. Guidance for AFEA Developers and Prospective Users

Guidance for AFEA developers	Guidance for prospective AFEA users
• **Don't overpromise:** The risks mentioned in this article in the context of a litigious society mean that developers should assume responsibility, not just for the safety of their products, but also for their efficacy. Be cautious in marketing, sales, and promotion. • **Invest in R&D that includes the "human element":** Develop interdisciplinarity in teams and in individual researchers. For instance, Facial Action Coding System (FACS) workshops can provide a basis for understanding facial behavior as well as a path to becoming a certified FACS coder. This is an entry point to the psychological, political, and societal concerns regarding AFEA's application. Focusing too much on one thing (e.g., algorithms or physiology) creates blind spots in the impacts of AFEA. • **Diversify your scientific team:** Scientific teams should employ experts in different fields who have diverse sociocultural and demographic backgrounds. This will foster awareness of how individual differences (gender, ethnicity, caste, class, nationality, geography, etc.) influence assumptions built into AFEA systems. This encourages sustainable growth, especially when all team members have a role in reducing the risk of systemic biases. Such teams should also embrace using FACS-certified human coders to audit AFEA coding (and vice versa) and mitigate the potential for systemic biases.	• **Don't buy into the hype:** Being a first mover carries significant risk. Early adopters often provide excellent case studies of "what not to do" for followers. AFEA's promise comes with risks and is not yet ready for high-profile or high-impact use. • **Stay current:** Technological developments can improve and extend capabilities, but often come with pitfalls. Just because a technology looks ready does not mean you are. Think carefully about the challenges discussed below and others that might arise. • **Proceed with caution:** Treat AFEA as emergent and apply it judiciously when failure is acceptable to you and your customer. AFEA can bring success, but there are no guarantees. Embed diversity, equity, and inclusion into each process to reduce the risk that AFEA implementation will reinforce systemic biases.

AFEA, however, is more complicated. Facial behavior [1] has been interpreted by humans and nonhuman primates for millennia (Darwin et al., 2002; Waller et al., 2020), and facial behavior is used to influence others every day, both consciously and subconsciously. Facial behaviors are often components of multimodal communications involving utterances (e.g., laughter, expostulation), posture (e.g., slumped shoulders in defeat), and hand and arm movements (e.g., pointing, supportive touches). Facial behaviors can sometimes even stand alone to transmit meaning: a knowing glance, the joy of a child's smile, the exasperation of an eye roll.

Fairly recently, researchers have examined (1) whether specific facial behavior correlates with various emotions and cognitions, and (2) when these behaviors are universal. A major pioneer of interpreting nonverbal behavior, Paul Ekman, built on existing scholarship to develop a system for codifying facial behavior: the Facial Action Coding System (FACS) (Ekman & Friesen, 1975).[2] The FACS involves 44 action units (AUs), each of which corresponds to a configuration of facial musculature, and 18 action descriptors (ADs), each of which involves movements of one or more facial muscles. FACS-certified coders

carry out step-by-step (using video, frame-by-frame) analyses, a laborious and time-consuming task (increasingly carried out by AFEA), to examine the occurrence, intensity, and timing of facial movements. FACS theory and findings have enabled coders to interpret sequences, peak intensities, and the onset/offset timing of AUs and ADs as emotions being signaled. For example, all smiles involve raising the lip corners using the zygomaticus muscle (AU 12), but different types of smiles convey different messages, revealed through the incorporation of other AUs. Such messages include amusement, affinity with another person, a dominance relationship, etc. (Rychlowska et al., 2017; Stewart et al., 2015).

AFEA is currently considered emergent. The current AFEA technology draws upon the FACS to identify AUs associated with specific emotional states. Most applications of AFEA are limited to algorithmically detected states of *basic* emotions: happiness, anger, fear, sadness, disgust, and surprise (Dupré et al., 2020). While several commercial AFEA technologies promise relatively fast learning curves,[3] open source software solutions could eventually offer greater customizability and algorithmic transparency. Still, there may be costs related to reduced accuracy,

---

[1] We use the term "behavior" rather than "expression" to refer to the configuration and movement of muscles in the face, thereby avoiding the fallacy of the converse, that a particular configuration or movement of facial muscles reflects (i.e., expresses) an internal state in a singular and interpretable way.

[2] Ekman inspired the television series *Lie to Me*, which featured a law enforcement and security consultant solving numerous difficult cases by analyzing non-verbal behaviors, including facial behaviors.

[3] We can discuss examples with those interested.

additional complexity, and uncertain support. AFEA applications can use video or images to detect AUs and ADs associated with specific emotions based on previously captured facial behavior or real-time analysis of live-streamed video. The software requires a mostly unobstructed, front-facing view with sufficiently high resolution to detect facial musculature, which is not always possible.

The performance of current commercial technologies is inferior to that of novice human raters. A recent study estimated a 54% correct assessment by AFEA software versus a 72% correct assessment by human raters; the best current AFEA systems recognize six basic emotions as accurately as human raters for "posed" behaviors but are significantly less accurate for spontaneous behaviors (Dupré et al., 2020). Low current AFEA accuracy, particularly for spontaneous behaviors, would argue against an at-scale, unsupervised application of AFEA. However, there are additional deeply rooted threats inherent to AFEA systems regarding challenges both *to* reliability and *of* reliability. Our use of the term *bias* in the following is clinical, referring to the predisposition toward (or against) some person or thing.

3 Challenges *to* Reliability

The primary challenges to reliability are found in four biases seen in AFEA output: simplicity bias, monomodal bias, environmental bias, and individual difference bias.

Simplicity bias refers to basic emotion theory's assumption of a one-to-one correspondence between facial behavior and emotion. In this expectation, the face provides a consistent readout of internal physiology. However, the richness of human emotional experience is not decomposable into the six basic emotions mentioned above (happiness, anger, fear, sadness, disgust, and surprise) because they do not account for emotions such as amusement, pride, love, disappointment, shame, exasperation, etc. (Barrett et al., 2019). Given the 44 AUs and 18 ADs identified by the FACS, facial behavior is obviously highly complex. AFEA software identifies, for example, the basic emotion of "happiness" indicated by the raised lip corners (zygomaticus muscle, AU 12) that are characteristic of all smiles, but does not consider a smile's social intent, signaled by the presence, intensity, and timing of other AUs and ADs (Rychlowska et al., 2017; Stewart et al., 2015).

Monomodal bias refers to reliance on facial behavior as sufficient to infer the emotional state and behavioral intent. Humans communicate through a combination of facial behaviors, verbal utterances, and other bodily movements. For example, a smile accompanied by raucous laughter and a shaking torso might indicate amusement, while a similar smile accompanied by a

relaxation of the body and a slowing heart rate might indicate contentment. Smiles can be interpreted differently depending on various physiological indicators. More measures may be needed for accurate interpretation.

Environmental bias refers to the inability to account for the external factors that predispose facial behaviors. For example, an antagonistic interlocutor (or even the perception of antagonism) can elicit different responses than a perceived friendly interlocutor. Therefore, the facial behaviors of different job candidates or crime suspects, for example, can be challenging to interpret. Similarly, a traveler with agoraphobia (fear of crowds) may display facial behavior at a busy airport security checkpoint that an AFEA system cannot distinguish from that of a traveler with forged documents.

Individual difference bias refers to the inability to account for genetic and life history factors that shape a subject's traits. Humans can do this innately: infants process the faces around them and, with experience, begin to understand the influence of facial behavior on others in social situations (Rosenberg & Ekman, 2020). This can unfold differently as people develop, depending on the biological, family, sociocultural, and economic factors in one's life history.

4 Challenges *of* Reliability

Reliable identification through computerized facial analysis might almost be a reality, but inherent biases have not been sufficiently solved to make AFEA reliable in the detection of emotions. Even if AFEA can be brought to the point of sufficient reliability, concerns about whether we *should* detect and act upon AFEA output will remain. The concerns here specifically relate to negativity bias, transparency bias, systemic bias, and subjectivity bias.

Negativity bias is the human tendency to focus more on negative than positive elements. Basic emotion theory characterizes only one of six basic emotions as affectively "positive" (happiness), four as "negative" (anger, fear, sadness, disgust), and one as neutral (surprise) (Ortony & Turner, 1990). Since humans have a tendency to focus on the negative (Baumeister et al., 2001), AFEA could encourage coercion and control as opposed to coordination and cooperation.

Transparency bias prioritizes data transparency over information privacy, which is often justified in terms of human progress and societal good. Artificial intelligence and digitalization enable data collection to infer thoughts and behavioral intentions (Zuboff, 2019). However, facial recognition technology has come under fire from privacy advocates, and the use of AFEA may be vulnerable to similar controversy in that it could impute internal thoughts or feelings from precognitive

or involuntary physiological responses to stimuli. Recognition of this threat was already expressed in a seminal work of modern US privacy thought from 1890:

> *The circumstance that a thought or emotion has been recorded in a permanent form renders its identification easier, and hence may be important from the point of view of evidence, but it has no significance as a matter of substantive right. If, then, the decisions indicate a general right to privacy for thoughts, emotions, and sensations, these should receive the same protection, whether expressed in writing, or in conduct, in conversation, in attitudes, or in* **facial expression**. (Brandeis & Warren, 1890, p. 206, emphasis added)

Systemic bias favors outcomes that preserve the characteristics of existing processes and structures, often to the detriment of marginalized groups. For example, computerized facial recognition has come under scrutiny for racial and gender biases (Spisak, 2022). The world's largest scientific computing society has recommended suspending the use of facial recognition systems by private and government organizations,[4] and a recent article in *Nature* has urged similar caution (Castelvecchi, 2020). Similar concerns have been voiced regarding machine learning algorithms that automate or augment decision-making about individuals. In one high-profile case, Amazon's hiring algorithm proved to be biased against women.[5] AFEA output could similarly be swept up in such controversy. Although AFEA systems are still emergent, they could be subject to similar risks or risks associated with combining multiple techniques, such as AFEA, facial recognition systems, and machine learning.

Subjectivity bias is the failure to recognize that what is morally right is subjective and tied to diverse sociocultural origins and value systems. This is difficult to address, even if structural inequities, privacy challenges, and negativity biases are fixed. If there is no universally accepted "good," and even high-stakes criminal justice cases are decided based on different beliefs about progress, fairness, justice, etc., it will be hard to make an AFEA system that works for everyone. Routine examples illustrate this difficulty: Should a large retail chain upgrading in-store security cameras with devices that capture customers' facial behavior exploit customer engagement and marketing opportunities? Should a teenage boy browsing the condom aisle and signaling fear through AFEA be sent a "push" notification? If so, should the push notification advise "Skip the checkout line and pay right from your phone," or offer an encouraging message like "You got

this!" (with a fist-bump emoji)? Or should the push notification ask a cautionary question such as: "Are you sure you're ready for this?" (perhaps with a link to additional information), or simply say nothing at all? Privacy issues aside, each option is morally contested.

5 Concluding Remarks

Facial recognition is currently being used and AFEA is evolving. Both are seeing rapid technological change, but such change carries risks as well as promise. Unattended risks can exacerbate previous problems and create new and unforeseen challenges. Facial recognition warrants caution but can likely be handled through normal risk assessment. However, this essay primarily focuses on the challenges posed by AFEA. Multiple factors limit AFEA's ability to reliably infer individuals' thoughts and feelings, and even if proven reliable, there remain concerns. AFEA could eventually be useful and beneficial for societies, organizations, and individuals, and we are optimistic over the long run. However, in this editorial, we offer pragmatic near-term guidance. Success in facial recognition does not foreshadow success with AFEA because judging human emotion or behavioral intention is much harder than simple identification. Organizations should be realistic in their expectations, cautious in their implementations, and critical when trying to predict potential negative impacts.

The advice above is aimed at organizations that are thinking about using AFEA or are developing it or experimenting with it now. From a public policy standpoint, AFEA is moving fast, meaning that policy observations from late 2022 might no longer apply in just a few years. Thus, the ongoing assessment of policy issues regarding AFEA, as well as facial recognition, is important. Given the advancing state of the art, regulatory bodies should avoid impeding innovation when benefits or costs are modest. For example, facial recognition to gain access to a device like a cell phone might fall into this category. But the trade-offs become more controversial when benefits and costs go up. For instance, a facial recognition system that is used to identify suspected criminals or grant access to an entitlement could precipitate concern among some stakeholders and lead to calls for prohibiting or (in rarer cases) requiring the technology's use.

However, the potential for controversy increases dramatically with AFEA. It presents a much more complicated realm of challenges than facial recognition, and the technology is not yet reliable enough to be routinely used to judge emotion or behavioral intent. We would not recommend that

---

[4] https://www.acm.org/binaries/content/assets/public-policy/ustpc-facial-recognition-tech-statement.pdf

[5] https://www.reuters.com/article/us-amazon-com-jobs-automation-insight/amazon-scraps-secret-ai-recruiting-tool-that-showed-bias-against-women-idUSKCN1MK08G

policy makers prohibit or require the use of AFEA, and we caution organizations against using AFEA to make decisions that impose costs or grant benefits to people. However, we encourage experimentation with the technology to determine how it might be used reliably, and how we, as a society, might wish it to be used once it becomes reliable.

Acknowledgements

References

Barrett, L. F., Adolphs, R., Marsella, S., Martinez, A. M., & Pollak, S. D. (2019). Emotional expressions reconsidered: Challenges to inferring emotion from human facial movements. *Psychological Science in the Public Interest*, *20*(1), 1-68.

Baumeister, R. F., Bratslavsky, E., Finkenauer, C., & Vohs, K. D. (2001). Bad is stronger than good. *Review of General Psychology*, *5*(4), 323-370.

Brandeis, L., & Warren, S. (1890). The right to privacy. *Harvard Law Review*, *4*(5), 193-220.

Castelvecchi, D. (2020). Is facial recognition too biased to be let loose? *Nature*, *587*(7834), 347-350.

Darwin, C., Ekman, P., & Prodger, P. (2002). *The expression of the emotions in man and animals*. Oxford University Press.

Dupré, D., Krumhuber, E. G., Küster, D., & McKeown, G. J. (2020). A performance comparison of eight commercially available automatic classifiers for facial affect recognition. *Plos One*, *15*(4), Article e0231968.

Ekman, P., & Friesen, W. V. (1975). *Unmasking the face: A guide to recognizing emotions from facial clues*. Prentice-Hall.

Ortony, A., & Turner, T. J. (1990). What's basic about basic emotions? *Psychological Review*, *97*(3), 315-331.

Rosenberg, E. L., & Ekman, P. (2020). *What the face reveals: Basic and applied studies of spontaneous expression using the facial action coding system (FACS)*. Oxford University Press.

Rychlowska, M., Jack, R. E., Garrod, O. G. B., Schyns, P. G., Martin, J. D., & Niedenthal, P. M. (2017). Functional smiles: Tools for love, sympathy, and war. *Psychological Science*, *28*(9), 1259-1270.

Spisak, B. R. (2022). Complex faces and naïve machines: A commentary on facial perceptions of age, gender, and leader preferences in the age of AI. *Politics and the Life Sciences*, *41*(1), 147-149.

Stewart, P. A., Bucy, E. P., & Mehu, M. (2015). Strengthening bonds and connecting with followers: A biobehavioral inventory of political smiles. *Politics & the Life Sciences*, *34*(1), 73-92.

Waller, B., Julle-Daniere, E., & Micheletta, J. (2020). Measuring the evolution of facial "expression" using multi-species FACS. *Neuroscience & Biobehavioral Reviews*, *113*, 1-11.

Zuboff, S. (2019). *The age of surveillance capitalism: The fight for a human future at the new frontier of power*. Profile books.

About the Authors

Jeffrey K. Mullins is an assistant professor of information systems in the Sam M. Walton College of Business at the University of Arkansas. His research areas include emotion, cognition, and ethics in information systems, and the IT-enabled convergence of work and play. His work has appeared or is forthcoming in *MIS Quarterly*, *Journal of Business Ethics*, *Journal of Business Research*, and other outlets. He has over a decade of IT experience at a Fortune 100 firm in areas including unstructured data management, IoT, e-commerce, and business analytics. He received his PhD from the University of Arkansas.

Patrick A. Stewart is a professor of political science and public administration at the University of Arkansas. He is a certified Facial Action Coding System (FACS) coder whose research concentrates on the emotional response to leaders including the observable audience response of applause, laughter, and booing during presidential debates. He has published over 75 books, articles, and chapters; his work on nonverbal communication and leadership has been featured in the *New York Times*, the *Washington Post*, and *The Conversation* and reported on in multiple popular media outlets. He received his PhD from Northern Illinois University.

Thomas J. Greitens is a professor of public administration at Central Michigan University. His research explores the administrative challenges of implementing innovation in the public sector, from online service delivery to technology adoption to privatization. His work has appeared in *Public Administration Review, Administration & Society, Public Performance & Management Review*, and in several books on public management. He received his bachelor's degree and MPA degree from Arkansas State University, and his PhD from Northern Illinois University.