Robots, Emotions, and the Lure of Complexity

FREDERIK VEJLIN

School of Communication and Culture, Department of Digital Design and Information Studies, Aarhus University

IN REVIEW

How That Robot Made Me Feel
Ericka Johnson
Cambridge, MA: MIT Press, 2025. 269 pp.

Ericka Johnson's edited volume, *How That Robot Made Me Feel*, takes up the well-worn question of whether robots or machines can have emotions from a somewhat unconventional angle. Rather than focusing strictly on robots' emotional capabilities, or lack thereof, the book reverses the question by asking how robots make *us* feel, the reader included. This sounds simple enough, but the book pulls this 'us' apart almost immediately, as if by design: users, researchers, designers, and publics all play different roles across the volume, and whose emotions are at stake, how they are measured, and by whom they are interpreted are issues that none of the chapters settle in quite the same way. As Johnson recounts in the introduction, the motivation for the book grew from a moment of productive misunderstanding at a seminar where she and a roboticist discovered they meant entirely different things by the term 'emotions.' For the roboticist, emotions were signals to be read through, for example, facial movements captured by a webcam and mapped onto discrete categories through software. For Johnson, emotions were entangled with context, power, culture, and lived experience, irreducible to any such mapping. The nine chapters that follow approach this gap from positions so varied (critical race studies, cognitive science, design fiction, Buddhist philosophy, among others) that the gap itself, and how it manifests in each, becomes, I think, the book's most interesting intellectual provocation.

The volume positions itself at the intersection of several traditions, drawing most prominently on feminist affect theory and science and technology studies to challenge what it presents as the dominant paradigm in human-robot interaction research and affective computing: the model of discrete, universal, individually embodied emotions that can be computationally read and classified, derived mainly from Paul Ekman's work on basic emotions. A model that has, as it happens, been challenged within psychology itself for some time (Barrett, 2017). This paradigm, formalised through Picard's (1997) foundational work on affective computing and operationalised through tools like the Facial Action Coding System, is the volume's recurring critical target (for an earlier critique, see Suchman, 2007). Against it, several contributors variously mobilise critical disability studies and critical race theory to expose the colonial and ableist genealogies of emotion recognition technology, and dramaturgical and labour theory to show the relational work that produces encounters with 'emotional' robots in practice. Others draw on enactivist and phenomenological approaches to embodiment and intentionality and, most ambitiously, Buddhist philosophy to propose alternative ontologies of emotion grounded in interdependence and ethical cultivation rather than discrete individual states. The range is deliberately heterogeneous, and the volume makes no attempt to synthesise these into a unified framework, staking its contribution instead on the claim that this very multiplicity demonstrates something about its object, namely that the complex configurations between emotions and machines are irreducible to any single paradigm as well as any simplified model. Whether this amounts to an argument or merely gestures toward one is a question the volume leaves open.

The volume's most forceful critical contribution comes from Jennifer Rhee, whose concept of emotion AI as 'imposition,' understood as the unilateral assignment of emotion categories that simultaneously strips people of authority over their own experience, provides a vocabulary that resonates through the collection even where it is not explicitly taken up. Drawing on critical autism studies and critical race theory, Rhee traces how emotion AI's foundational commitments to the theory of basic emotions and the theory of mindblindness reproduce patterns of deauthorisation along lines of race and disability. This analysis anchors the volume's broader argument that the politics of emotion in robotics are simultaneously technical, structural, and historical (see also Atanasoski and Vora, 2019; Crawford, 2021; Rhee, 2018; Srnicek, 2025).

At the other end of the collection, Daniel White and Hirofumi Katsuno's chapter on Buddhist models of emotion and their application in Japanese robotics,

from Mori Masahiro's Buddhist-inflected engineering philosophy to the Android Kannon Mindar at Kōdaiji temple in Kyoto, offers perhaps the most unexpected contribution, proposing the concept of 'mindful machines' to capture the idea that human-robot relations might be understood as sites of mutual ethical cultivation rather than instrumental extraction or one-sided imposition. Read alongside the broader critical research on Japanese robotics culture (Robertson, 2007, 2014, 2018; see also, e.g., de Togni, 2024; Frumer, 2018; Kovacic, 2018; Šabanović, 2014; Sone, 2017), the chapter raises productive questions about how engineering, national techno-imaginaries, religious practices, and philosophical tradition intersect in ways that resist easy generalisation.

Between these poles, the volume's empirical and experimental chapters demonstrate the relational complexity of what happens when robots and emotions are entangled in practice. Ethnographic work in Swedish dementia care homes reveals the sophisticated emotional labour caregivers perform to make robot animals 'work,' while a design fiction about a haptic, nonhumanoid AI agent explores what a 'feel first, reason later' design philosophy might look like. Though one might note that the insight that robots require rather than replace relational care work is by now fairly well established in the care robotics literature (e.g., Chevallier, 2023; Jeon et al., 2020; Lipp, 2023, 2024; Šabanović and Chang, 2016; Shin and Jeon, 2024; Wright, 2018, 2023). It is also worth noting that the labour of making a simplified or unpredictable entity workable within a complex institutional setting is not unique to robots, and that attending to what this produces, rather than mainly what it compensates for, might open different analytical paths attuned to the forms of expertise, affective repertoires, and institutional arrangements that emerge around robots in practice. Some chapters come closer to this than others.

Even so, the volume's contribution lies less in any single finding than in the cumulative effect of juxtaposing such different approaches to the same question. Together, the chapters amount to something like a claim about 'complexification,' that is, about thickening our understanding of something that dominant frameworks in robotics tend to flatten and universalise. As an anthropologist of technoscience researching social robotics and artificial intelligence in Japan, I find this impulse sympathetic but also in need of its own examination. The risk of 'humanistic critique' of robotics, especially when it positions itself as bringing nuance, context, and power to a field imagined as naïve about these things, is that it can reproduce its own form of narrowing. In such cases, robotics is easily framed as a monolith of quantification and universalism, to which the social sciences and humanities bring the gift of critique and complexity (Vejlin, 2024; see also Dan-Cohen, 2017, 2021; Hoeck, 2025; Mol and Law, 2002).

Several chapters come close to this, particularly when they invoke Ekman and Picard as shorthand for everything wrong with how emotions are handled in HRI. But the perceived thinness of technical practice may tell us as much about the critical frame being deployed as about the practices themselves. Robotics and HRI are also diverse, contested, and internally reflexive fields, though perhaps not in the ways we are familiar with in the humanities and social sciences. In Japan, as White and Katsuno's own chapter demonstrates, the relationship between engineering practice, sociocultural imaginaries, and emotional design has always been more entangled than the 'basic emotions' label would suggest (see also Hoeck, 2024; Katsuno, 2011; White and Katsuno, 2021). A critical engagement through complexification and multiplicity only really works, I think, if its object is actually as monolithic as the critique assumes, and, on my reading, the volume's own best chapters quite clearly show us that it is not.

This is not to say that the volume's critiques are misplaced; Rhee's analysis of deauthorisation, for instance, is a necessary intervention. But the most productive contributions are those that also find complexity, ambiguity, and indeterminacy within the practices of robotics itself rather than opposing 'rich' humanistic analysis to 'thin' technical practice.

When caregivers construct the emotional roles of robot animals through their own skilled labour, or when Japanese roboticists are shown to have long grappled with questions of consciousness and compassion as part of their engineering practice, the distinction between complexification and simplification becomes murky, and the analysis, in turn, becomes more interesting for it.

The question of whose complexity counts is also, of course, a question about whose knowledge counts, and it is here that the volume's relevance for readers of this journal is most obvious. The feminised labour of care, the intersections of gender, race, and disability in emotion AI, the gendered bodies of humanoid robots, and, most broadly, the question of epistemic authority over emotion – who gets to define it, measure it, and speak about it – run through the collection even where gender is not always explicitly foregrounded (see also, e.g., Adam, 1998; Weber, 2005; Perugia and Lisy, 2023; Robertson, 2010). The volume's methodological diversity can itself be read as a feminist intervention, asserting the legitimacy of introspection, design fiction, and ethnographic observation alongside the quantified experimental paradigm that, to an extent, still dominates HRI. But there is also a risk that assembling diverse critical perspectives around a shared target might come to substitute for the harder work of developing claims that are not already available within each contributor's own tradition. A concluding chapter drawing these threads together might have helped. As it stands, the volume's most important insight, namely that the politics of emotional entanglement with robots cannot be settled from within any single framework, emerges from the juxtaposition of chapters rather than from any one argument. But for an interdisciplinary conversation that is still in the midst of defining its terms of engagement, the irresolution may itself be the point.

LITERATURE

Adam, A. (1998). *Artificial Knowing: Gender and the Thinking Machine*. Routledge.

Atanasoski, N., & Vora, K. (2019). *Surrogate Humanity: Race, Robots, and the Politics of Technological Futures*. Duke University Press. https://doi.org/10.1215/9781478004455

Barrett, L. F. (2017). *How Emotions are Made: The Secret Life of the Brain*. Houghton Mifflin Harcourt.

Chevallier, M. (2023). Staging Paro: The care of making robot(s) care. *Social Studies of Science*, *53*(5), 635–659. https://doi.org/10.1177/03063127231178520

Crawford, K. (2021). *Atlas of AI: Power, Politics, and the Planetary Costs of Artificial Intelligence*. Yale University Press. https://doi.org/10.2307/j.ctv1ghv45t

Dan-Cohen, T. (2017). Epistemic artefacts: on the uses of complexity in anthropology. *Journal of the Royal Anthropological Institute*, *23*(2), 285-301. https://doi.org/10.1111/1467-9655.12607

Dan-Cohen, T. (2021). *A Simpler Life: Synthetic Biological Experiments*. Cornell University Press. https://doi.org/10.7591/cornell/9781501753442.001.0001

De Togni, G. (2024). Staging the robot: Performing techno-politics of innovation for care robotics in Japan. *East Asian Science, Technology and Society: An International Journal*, 196-213. https://doi.org/10.1080/18752160.2023.2295144

Frumer, Y. (2018). Cognition and emotions in Japanese humanoid robotics. *History and Technology*, *34*(2), 157-183. https://doi.org/10.1080/07341512.2018.1516849

Hoeck, K. (2024). Android Existence: The Affect of Artificial Vitality. *HAU: Journal of Ethnographic Theory*, *14*(3), 733-747. https://doi.org/10.1086/732626

Hoeck, K. (2025). Technological Totemism: Gendered Typification as Problematic Model for Robotic Existence. *Current Anthropology*, *66*(5), 759-772. https://doi.org/10.1086/737430

Jeon, C., Shin, H., Kim, S., & Jeong, H. (2020). Talking over the robot: A field study of strained collaboration in a dementia-prevention robot class. *Interaction Studies*, *21*(1), 85–110. https://doi.org/10.1075/is.18054.jeo

Katsuno, H. (2011). The robot's heart: Tinkering with humanity and intimacy in robot-building. *Japanese Studies*, *31*(1), 93–109. https://doi.org/10.1080/10371397.2011.560259

Kovacic, M. (2018). The making of national robot history in Japan: monozukuri, enculturation and cultural lineage of robots. *Critical Asian Studies*, *50*(4), 572-590. https://doi.org/10.1080/14672715.2018.1512003

Lipp, B. (2023). Caring for robots: How care comes to matter in human-machine interfacing. *Social Studies of Science*, *53*(5), 660–685. https://doi.org/10.1177/03063127231175630

Lipp, B. (2024). Robot drama: Investigating frictions between vision and demonstration in care robotics. *Science, Technology, & Human Values*, *49*(2), 318–343. https://doi.org/10.1177/01622439221131657

Mol, A., & Law, J. (2002). Complexities: An introduction. In J. Law & A. Mol (Eds.), *Complexities: Social Studies of Knowledge Practices* (pp. 1–22). Duke University Press.

Perugia, G., & Lisy, D. (2023). Robot's gendering trouble: a scoping review of gendering humanoid robots and its effects on HRI. *International Journal of Social Robotics*, *15*(11), 1725-1753. https://doi.org/10.1007/s12369-023-01038-x

Picard, R. (1997). *Affective Computing*. MIT Press.

Rhee, J. (2018). *The Robotic Imaginary: The Human and the Price of Dehumanized Labor*. University of Minnesota Press.

Robertson, J. (2007). Robo sapiens japanicus: Humanoid robots and the posthuman family. *Critical Asian Studies*, *39*(3), 369–398. https://doi.org/10.1080/14672710701527378

Robertson, J. (2010). Gendering humanoid robots: Robo-sexism in Japan. *Body & Society*, *16*(2), 1–36. https://doi.org/10.1177/1357034X10364767

Robertson, J. (2014). Human rights vs. robot rights: Forecasts from Japan. *Critical Asian Studies*, *46*(4), 571–598. https://doi.org/10.1080/14672715.2014.960707

Robertson, J. (2018). *Robo Sapiens Japanicus: Robots, Gender, Family, and the Japanese nation*. University of California Press. https://doi.org/10.1525/california/9780520283190.001.0001

Šabanović, S. (2014). Inventing Japan's 'robotics culture': The repeated assembly of science, technology, and culture in social robotics. *Social Studies of Science*, *44*(3), 342-367. https://doi.org/10.1177/0306312713509704

Šabanović, S., & Chang, W.-L. (2016). Socializing robots: Constructing robotic sociality in the design and use of the assistive robot PARO. *AI & Society*, *31*, 537–551. https://doi.org/10.1007/s00146-015-0636-y

Shin, H., & Jeon, C. (2024). The robotic multi-care network: A field study of a "robot grandchild" in South Korea. *East Asian Science, Technology and Society: An International Journal*, *18*(2), 177–195. https://doi.org/10.1080/18752160.2024.2348304

Sone, Y. (2017). *Japanese Robot Culture: Performance, Imagination, Modernity*. Palgrave Macmillan. https://doi.org/10.1057/978-1-137-52527-7

Srnicek, N. (2025). *Silicon Empires: The Fight for the Future of AI*. Polity.

Suchman, L. (2007). *Human-Machine Reconfigurations: Plans and Situated Actions* (2nd ed.). Cambridge University Press. https://doi.org/10.1017/CBO9780511808418

Vejlin, F. (2024). Kritiske komplikationer: Kompleksitet og etik i antropologisk forskning. In J. Kofoed & N. J. Koefoed (Eds.), *Etiske øjeblikke i humanistisk forskning* (pp. 231–258). Aarhus Universitetsforlag.

Weber, J. (2005). Helpless machines and true loving care givers: A feminist critique of recent trends in human-robot interaction. *Journal of Information, Communication and Ethics in Society*, *3*(4), 209-218. https://doi.org/10.1108/14779960580000274

White, D., & Katsuno, H. (2021). Toward an affective sense of life: Artificial intelligence, animacy, and amusement at a robot pet memorial service in Japan. *Cultural Anthropology*, *36*(2), 222–251. https://doi.org/10.14506/ca36.2.03

Wright, J. (2018). Tactile care, mechanical hugs: Japanese caregivers and robotic lifting devices. *Asian Anthropology*, *17*(1), 24-39. https://doi.org/10.1080/1683478X.2017.1406569

Wright, J. (2023). *Robots Won't Save Japan: An Ethnography of Eldercare Automation*. Cornell University Press. https://doi.org/10.7591/cornell/9781501768040.001.0001