Cultural Anthropology for Social Emotion Modeling: Principles of Application toward Diversified Social Signal Processing

Daniel White
Freie Universität Berlin
Berlin, Germany
daniel.white@fu-berlin.de

Hirofumi Katsuno
Doshisha University
Kyoto, Japan
hkatsuno@mail.doshisha.ac.jp

***Abstract*—The practice of modeling social emotions has benefited from interdisciplinary engagements with other fields in the hard and human sciences; however, perspectives from cultural and social anthropology have been limited. This has at times resulted in the integration of emotion theories into emotion modeling that emphasize the universal communicability of social signals of emotion at the expense of accounting for cultural diversity evidenced in the ethnographic record. This paper outlines methods and findings of a collaborative effort between cultural anthropologists and engineers to create platforms for interdisciplinary communication and emotion modeling practices more sensitive to cultural diversity and better protected from risks of ethnic, racial, and ethnocentric bias. The paper presents five principles for applying anthropological perspectives to emotion modeling and ultimately argues for a consideration of design strategies for social signal processing based on recent ethnographic evidence of evolving human-robot relationships in Japan.**

Keywords—affect, affective robotics, cultural anthropology, emotion, multispecies society, social emotion, social signal processing, thick description

I. Introduction

While research on emotion signaling processes within social robotics has increasingly integrated interdisciplinary perspectives from fields such as biology, psychology, and cognitive neuroscience, views from cultural and social anthropology have been underrepresented. Given the vast amount of data on the cultural variability of emotional expression in the ethnographic record, the lack of attention to such diversity can limit the flexibility and adaptability of emotion-recognition systems, such as in social and companion robots designed to connect affectively with human users. This has at times resulted in the overreliance on universal theories of emotion communication in the design and implementation of emotion models for artificial social agents. Complicating matters further is the risk associated with implementing robots in experimental settings that may generate data on human-robot interaction (HRI) susceptible to ethnic and racial bias, gender discrimination, and the ethnocentric exclusion of exceptional yet important results due to the limited consideration of parameters within social signal processing (SSP) [1, 2]. Given that both the burden of responsibility and benefits of reciprocity for cross-disciplinary dialogue may be shared equally by researchers from both anthropology and social robotics, with implications well beyond these fields, this paper outlines efforts and findings of a recent interdisciplinary collaboration to cultivate more culturally-sensitive design practices for SSP within affective robotics. It further proposes principles of application for diversified, data-rich, interactive, and culturally-sensitive practices for modeling social emotions more generally.

To provide context and supporting evidence for these principles, the paper draws on the anthropological background of the authors and on ethnographic evidence collected through interviews and interaction with engineers in robotics labs in Japan. It organizes its findings through the following sections. Section II of the paper outlines some primary challenges facing social emotion modeling. Section III then describes the methods the researchers used to address some of these challenges, including the creation of interdisciplinary workshops as platforms for generating more culturally-sensitive and data-rich emotion modeling practices. Section IV outlines five preliminary principles for applying anthropological perspectives to SSP and emotion modeling. Section V discusses the findings of collaborative fieldwork between anthropologists and engineers and explains the paper's central argument. Stated summarily, it proposes that given human affect is constructed interactively and variably in coordination with other human and non-human agents, robot designers interested in SSP may benefit from strategizing how to invite the pleasurable adaptation of humans to machines in addition to adapting machines to humans. The paper further suggests that treating *all* emotions as fundamentally social and interactive affords a cognitive shift conducive to more dynamic emotion modeling.

II. Overview of Challenges

A classic example from cultural anthropology highlights a major challenge for evaluating social signals. Drawing on the work of philosopher Gilbert Ryle, the cultural anthropologist Clifford Geertz [3, pp. 6-7] asks his readers to consider the polysemic complexity of a wink in a section that is worth citing at length:

Consider…two boys rapidly contracting the eyelids of their right eyes. In one, this is an involuntary twitch; in the other, a

978-1-7281-3891-6/19/$31.00 ©2019 IEEE

conspiratorial signal to a friend. The two movements are, as movements, identical; from an l-am-a-camera, "phenomenalistic" observation of them alone, one could not tell which was twitch and which was wink, or indeed whether both or either was twitch or wink. Yet the difference, however unphotographable, between a twitch and a wink is vast; as anyone unfortunate enough to have had the first taken for the second knows. The winker is communicating, and indeed communicating in a quite precise and special way: (1) deliberately, (2) to someone in particular, (3) to impart a particular message, (4) according to a socially established code, and (5) without cognizance of the rest of the company. As Ryle points out, the winker has not done two things, contracted his eyelids and winked, while the twitcher has done only one, contracted his eyelids. Contracting your eyelids on purpose when there exists a public code in which so doing counts as a conspiratorial signal is *winking…*

That, however, is just the beginning. Suppose, he continues, there is a third boy, who, "to give malicious amusement to his cronies," parodies the first boy`s wink, as amateurish, clumsy, obvious, and so on. He, of course, does this in the same way the second boy winked and the first twitched: by contracting his right eyelids. Only this boy is neither winking nor twitching, he is parodying someone else's, as he takes it, laughable, attempt at winking…One can go further: uncertain of his mimicking abilities, the would-be satirist may practice at home before the mirror, in which case he is not twitching, winking, or parodying, but rehearsing…Complexities are possible, if not practically without end, at least logically so. The original winker might, for example, actually have been fake-winking, say, to mislead outsiders into imagining there was a conspiracy afoot when there in fact was not, in which case our descriptions of what the parodist is parodying and the rehearser rehearsing of course shift accordingly. But the point is that between what Ryle calls the "thin description" of what the rehearser (parodist, winker, twitcher . . .) is doing ("rapidly contracting his right eyelids") and the "thick description" of what he is doing ("practicing a burlesque of a friend faking a wink to deceive an innocent into thinking a conspiracy is in motion") lies the object of ethnography.

Geertz's articulation of Ryle's wink has long served as exemplary of what kind of data anthropologists are seeking when it comes to understanding cultural variability. This "thick description" of events captures the richness and complexity of social interactions. And in Geertz's literary description of the nuances of social signaling, he artfully leverages qualitative description to communicating data on cultural complexity.

While the largely quantitative methods of emotion modeling in engineering differ from the mostly qualitative ones of emotion description in anthropology, much can be learned about what each method affords the other. Consider one particular challenge highlighted by computer scientists Aylett and Paiva [4, p. 253]:

In order to implement any model on a computer, the model itself must be sufficiently specific. From this perspective, many psychological models are not usable as they stand, but must be operationalized. Qualitative relationships must be quantified…Thus, when computer scientists select models from psychology, they tend to favour those that are already sufficiently specific or that can be made so relatively easily.

This qualitative-quantitative gap in methodology is what in part divides anthropology from engineering, leaving many engineers to draw from those psychological models of emotion most amenable to programing in code. And yet, although the language and methodological approaches between disciplines may differ, the challenge of understanding social signals for researchers in affective robotics, as highlighted by Geertz, is similar to those in cultural anthropology: how to build an understanding of social complexity into models and theories of social interaction. The next section describes methods we apply in order to identify the most common challenges for implementing considerations of cultural variability in social emotion modeling.

III. METHODS

The authors of this paper, trained in cultural anthropology and having conducted fieldwork on human-robot relations in Japan, applied three primary methods to research. First, they conducted textual and discursive analysis of recent scientific literature on emotion modeling and social signaling in both English and Japanese. Second, they conducted ethnographic fieldwork in sites of robot design and human-robot interaction in Japan, including discussions with and observations of robotics engineers in labs, interviews with users and fans of companion robots such as AIBO and LOVOT, and participant observation in conferences, symposiums, and exhibitions focused on social robots. Third, they designed and ran preliminary trials of collaborative workshops between anthropologists and engineers sharing methodological perspectives on emotion modeling and social signaling. Given that the third method of data collection, collaborative workshops, can also serve as a reproducible platform for integrating considerations of cultural diversity in practices of emotion modeling, it is worth explaining their aims and structure in further detail.

The workshops we employ are titled Anthropology for Affective Robotics (AfAR) and are described as a platform for building collaboration between anthropologists and engineers working in robotics, artificial intelligence, and informatics. Workshops are hosted by the authors, who explain their aims to participants as the following: 1) create platforms for mutual exchange and benefit between anthropologists and engineers on the latest methods for modeling culture and emotion; 2) increase awareness of cultural variability in order to improve modeling practices in laboratories and companies working on robotics and AI; 3) establish sustainable connections between the human and hard sciences for thinking about the social, ethical, political, and legal implications of AI and robotics research.

Workshops are offered in either English or Japanese and consist of three primary parts: 1) description of individual and team research projects by lab members, with a focus on their approaches to or problems encountered with modeling culture and/or emotion; 2) introduction of basic principles of culture and emotion theory in anthropology, including latest research findings; and 3) a dialogue session including a) suggestions from workshop hosts for individual and group projects and b) suggestions from lab members for workshop hosts on the anthropological approach to culture and emotion. The workshops are in ideal cases filmed and content shared among all members, allowing for follow-up sessions and subsequent evaluations and feedback.

IV. FINDINGS

This section identifies some of the primary differences between engineering and anthropological approaches to social emotion and frames some principles for applying

anthropological perspectives to future practices of emotion modeling for artificial agents. Each principle is framed as a response to assumptions we identified as most common in engineering labs and emotion modeling literature; most distinct from contemporary ethnographic approaches to affect and emotion; and most limiting for designing SSP capacities more sensitive to cultural diversity and the interactive contexts by which emotional communication is mutually constructed within dynamic social environments.

A. Cultural Anthropological Principles on Social Emotions

1) Social emotions are culturally variable in both their expression and their physiological states. A common assumption guiding emotion modeling and SSP is that social emotions are rooted in paradigms of "basic" [5–8] or "primary" [9] emotions which can be identified in universally similar signals such as facial expression. However, as shown by anthropological work on the cross-cultural expression of emotion [10–15], emotional expressions are highly variable linguistically, as well as in bodily gestures, facial expressions, and other signals communicable to humans. Moreover, as further illustrated by work on affect in the social sciences [16], affective neuroscience [17], and social psychology [18], the physiological states triggering such signals cannot be identified in universal emotion circuits in the brain, or so-called "neurological fingerprints" [18], that correspond with basic discursive categories such as anger, fear, or joy. Rather, emotional states are just as much variously "constructed" [18] by nervous systems interacting with particular cultural environments as are the set of conventions of emotional expression and behavior characteristic of those environments. In other words, not only are emotional *expressions* socially variable but so too are emotional *states,* given they are cultivated through processes of social of somatic development over time. We distinguish between these discursive and somatic aspects of emotion by calling the first *emotion* and the second *affect*, while also drawing attention to how these components themselves are conditioned interactively in coordination with other agents, human or otherwise [16, 19].

The consequence of this finding for building more complex SSP paradigms, and a point increasingly identified by recent work in human-robot interaction [20–22], is that researchers must take care in disambiguating affective states from emotional expressions and behavior, given they cannot be uniformly and universally correlated. For example, while facial recognition coding systems such as FACS [23] and adaptations in toolkits such as OpenFace [24] thus offer powerful means for generating data on emotion expressions, the interpretation of such data should be exercised in consideration of how affect, emotion, and behavior are grounded in interactive processes of adjustment, negotiation, and coordination in local contexts and conditions [22].

2) Social emotions are products of heterogenous cultural arrangements. A common misconception about culture is that one can draw clear boundaries around it and presume its members share similar dispositions fixed to a single parameter. While this assumption may afford high degrees of statistical probability for some parameters, such as associating an "American" individual with one that is also a "native speaker of English," this is of course not always the case. Contradictory cases highlight the consequences of conflating nationality with cultural identity in experimental design (i.e. positing "American" participants against "Japanese" ones as indicators of cultural difference). That a high degree of probability for this fact can be established in data sets collected from experimental settings that nonetheless divide groups by nationality in order to procure "cultural" data does not often lead to building emotion models that are more culturally diverse. Rather, dividing participants purely by nationality risks normalizing a view within emotion modeling that assumes cultural and national parameters are synonymous—a view that is at odds with ethnographic findings.

Alternative approaches sensitive to cultural heterogeneity would apply extensive qualitative inquiries on participant backgrounds and personal histories prior to experiments incorporating SSP and provide opportunities for meta-level reporting from participants on their experience and the data collected [25]. Assembling culturally and gender-diverse research teams as well as incorporating social scientists in decisions on setting parameters for training datasets in machine learning systems offer increased protections against designing overly-homogenous social signaling systems.

3) Social emotions are indexical. Another common assumption we identify in experimental designs, and make a point to contest, is that social emotions can be measured independently of context. As illustrated in the above example, the fact that high statistical probability can with some parameters be confirmed in groups categorized by nationality does not protect against exceptions resulting from context-contingent factors. When Ekman [26] and Friesen [27] found that Japanese individuals expressed disgust when watching violent videos just as American ones do, except when in the company of an authority figure, in which case they smile, they were articulating a social signal that is context-dependent: a "display rule" [28]. However, although this finding was critically important, and display rules well recognized by social signal researchers since [2], it did not nearly account for the degree of contextual and multimodal complexity—the content of the videos, the personal history of participants, their blood-sugar level, the design of the lab, the weather, the subtle smell of tobacco on the researcher's vest which reminded the participant of the habit she at long last recently broke—required for robust social signal modeling. Calling affective experiences "indexical" [16] means that the semantic content of similar social signals, their environmental triggers, and, importantly, the affective states they are associated with shift based on a multitude of factors particular to each person. While this is often very clear to researchers, the strategy of constructing stable laboratory and experimental conditions to limit environmental variables does not address the variability in individuals' personal histories.

Designing and building ample funding into project research designs that afford HRI experiments with a group of participants over the entire course of the project term allows for increased testing under different contexts at different times. Integrating training modules for both participants and researchers on specific aspects of introspection and emotional intelligence relative to the social signals featured in projects can also increase the

accuracy of self-reporting data. Further, treating individuals who join experiments as mere "participants" whose roles are confined to the production of quantitative data limits what kinds of data they can share. Engaging with them instead as individuals, collaborators, or interlocutors whose reactions and input can affect experimental settings at all stages of interaction expands the parameters for collecting their feedback, even if this can only be accounted for in the qualitative discussion section of research papers.

4) Social emotions are dynamic. A corollary of semantic indexicality is temporal mutability. A fundamental principle of cultural arrangements and the emotional patterns they condition is that they are open and responsive to change [29]. This contests common assumptions constructed within experimental settings that posit social emotions as rooted in stable cultural arrangements that are translatable from laboratory to organic settings. Although different groups of people may hold certain patterns of behavior or "structures of feeling" [30] in common, these are constantly reproduced on multiple temporal scales (historically, generationally, daily, even in moment-to-moment interaction). Moreover, these cultural changes are open to global processes that condition capacities to "affect and be affected" [31, 32] according to shifting constellations of capital, media, and information. This raises the stakes of the point stated in the first principle above that not only are social *expressions* time dependent, indexed by shifting arrangements of sign systems that mutate and evolve, but so too are the physiological states—or *affect*—those sign systems condition. Acknowledging this temporal variability requires close attention to how *discourse* (the way we talk about emotion) and *affect* (the way it becomes embodied) operate in an ongoing feedback loop with one another [19]. It further requires constant self-reminding that affective states can never be confidently correlated with social signals of their expression without referencing meta-reflections from interlocutors, even if those too are imperfect.

Emotion modeling strategies sensitive to the complexity of temporal mutability, as well as to how social signaling differs dramatically from lab to "real-world" settings, would seek creative means for thick data collection and mining outside labs, such as by drawing on methods of participant observation and long-term self-assessment. It would also avoid universalizing popular paradigms of robotics design, such as that of the "uncanny valley" [33], that do not incorporate a dimension measuring how human agents adjust affectively to robots through processes of interaction and accommodation over time [34, pp. 159-160]

5) Social emotions are mediated within interactive relations among both human and non-human arrangements. Although many models of social emotion ground emotion in interaction that is anthropocentric, such a perspective limits cataloguing how emotion and affect are generated in human relations with objects, environments, and rapidly mutating networks of material and digital culture. In his extensive philosophical reflection on affect, Baruch Spinoza observed that "No one has yet determined what the body can do…For no one has yet come to know the structure of the body so accurately that he could explain all its functions" [32, pp. 71-72]. While a positivist perspective may hold out hope that such a project could be completed with the right analytical tools of observation, or technologies of intervention, a cultural anthropological view argues that this project is in fact open and unending given that the body's capacities are attenuated, amplified, or otherwise affected by human relationships as well as by material culture. Material culture can include technologies such as social and companion robots that alter our affective states through interaction at the very moment they attempt to read and register it. This entangling of the semantic content of social signals and the somatic states they both signal and newly engender challenges researchers who model social emotions to move beyond an approach that divides a so-called *natural* category of "biologically processed social signals" that is "largely culturally invariant" from a *cultural* category that is socially constructed and diverse [35, p. 12].

An alternative perspective from cultural anthropology, supported by recent work in neuroscience [18, 36], would attend to the dynamic co-dependence of nature-culture and seek ways to understand how culture conditions the body even at the level of the autonomic nervous system. Applying this perspective to understanding social signaling within HRI would incorporate strategies to account for the ongoing aesthetic and interactive effects of human-robot design by tracing entire networks of relatability and affect activation. Perspectives drawn from Actor-Network Theory [37, 38] and mediation theory [39] offer methods for tracing and accounting for distributions of agency in systems that are not limited to humans, environmental factors, or technological objects as distinct entities but in fact exist as hybrid and composite figures capable of agency and even social signaling functions of their own.

V. Discussion

Although these five principles are in no way exhaustive, they offer some initial examples of how considerations of cultural diversity found in the ethnographic record and ongoing fieldwork can inform design practices for modeling social emotions. That social emotions cultivated within different cultural arrangements are variable, heterogenous, indexical, dynamic, and relationally mediated by human and non-human actors indicates not only challenges for SSP but also opportunities for new technological methods of exploring and accounting for the complexity of social signals even beyond what ethnographic methods of thick description can afford. While the above principles and suggestions do not serve as means to neatly bridge an obvious qualitative-quantitative gap in disciplinary approaches, they may nonetheless generate progress by stimulating creative communication across it.

One principal implication of our ethnographic findings that is perhaps unexpected from a perspective rooted in a study of *anthropos*, or "the human," is the degree to which determining how social signaling works among humans requires accounting for how it is mediated by artificial objects and agents. Evidence from past ethnographic accounts of human-robot relations [34, 40–42] as well as from our own observations of human-robot interactions in Japan shows that people can adapt to machines as much as machines can be programmed to adapt to people. This fact suggests the value of rethinking dominant paradigms in robotics in order to improve design practices for modeling

social emotion. For example, researchers following the synthetic model for emotion modeling have made important advances by patterning SSP after human behavior. One reason for this, articulated by Tsiourti [43, p. 2], comes from the evaluation that the "affect-expression capability of humans can serve as 'the gold standard' and a guide for defining design recommendations for multimodal expression of human-like affective states." We think Tsiourti's claim is correct. We also think that it could be added to, improved, and updated by drawing further on anthropological perspectives on HRI.

Anthropological and sociological evidence suggests that human affective states are mediated by the way technological objects are arranged within human and non-human networks [37, 38]. Given that forms of this mediation can result in new ways of expressing empathy, care, and affection that may be intelligible to a robot but not immediately to a human, we propose that designers thinking about social emotion within human-robot interaction may benefit from experimenting with emotion modeling platforms that invite humans to adjust to machines rather than uniformly adapting machines to humans.

For example, Groove X's robot LOVOT [44] incorporates a function whereby a human pressing the robot's nose will evoke a response somewhere between an irritated sneeze and a giggle. This functionality particular to LOVOT invites the human to "signal" affection (or a desire for reciprocal affection) in a way that is particular to the human-LOVOT relation. We have observed that such a function is easy and enjoyable for humans to learn as well as for LOVOT's software to record and encode with a clear meaning. Moreover, with the robot's facial recognition and tactile sensors, it also importantly affords multimodal signaling recognition in a real-world setting [45].

While the suggestion to increase efforts on building robotic platforms like the one described above in order to invite human accommodation to machines may not initially be viewed as classically anthropological, nor sympathetic to anthropocentric concerns within HRI, we argue that recent theoretical perspectives on multispecies societies [46–48] coupled with our own observations of human-robot relations in Japan illustrate how culturally-particular values on wellbeing are being reimagined in relation to artificial agents. These relationships can cultivate new affective capacities and potentially even new strategies for SSP that are sensitive to the needs and desires of human-robot relationships practiced in different cultural settings.

Finally, the point increasingly emphasized in HRI studies [20–22] that emotional expressions, behavior, and affective states are constructed interactively with one another leads us to submit a final proposition on the concept of "social emotions" that could only be articulated after outlining the above principles. In fact, we find some redundancy in the term "social emotions." From an anthropological perspective, affective states, behaviors, and emotional patterns of expression are constructed interactively within environments that are always already social. As Paul Dumouchel [49, pp. 1-2] has argued, "Emotions are social in that they are not the means but the *state* or the *being* of humans living together. The fact that we have an affective life is not a cause of but *is* the fact that, as beings, we are not completely independent of each other." Placed in the context of SSP for human-robot interaction, we can say that *all* emotions, and indeed all moments of experience more generally—even those passed in somatic and semantic privacy—incorporate an interactive and resolutely social component of conditioning; it is simply that one may need to integrate the agencies of environments, objects, and indeed increasingly of social robots to best account for it.

VI. Conclusion

Drawing on anthropological perspectives that illustrate the flexibility of emotion across human experience, this paper has aimed to initiate a platform for collaboration between anthropologists interested in analyzing the social complexity of emotion and engineers interested in modeling that complexity in artificially intelligent and emotionally-sensitive agents. Although the paper has illustrated some differences in approaches to theorizing and modeling emotion, it had done so not to criticize but to cultivate further collaboration in designing improved strategies of emotion modeling that draw on methodologies from both disciplines. We hope this collaboration may consequently augment capacities to account for cultural diversity while attenuating risks of ethnocentric bias not only in engineering but also in the humanities and human sciences. While we propose that one strategy for achieving this may be to leverage humans' affective adaptability in order to create partnerships with robots, we also emphasize that this purpose is not to limit human expressions of emotion. Rather, integrating cultural diversity into SSP allows researchers to more accurately observe and trace the already-expanding affective capacities of humans into more capacious and sustainable platforms for wellbeing in our multispecies-populated societies.

Acknowledgment

The authors thank three reviewers for critical insights into how to better navigate the interdisciplinary terrain of this paper. The project Model Emotion receives institutional support from Freie Universität Berlin and Doshisha University. Its additional contributors and outputs can be explored further at www.modelemotion.org.

References

[1] A. Vinciarelli, M. Pantic, and H. Bourlard, "Social signal processing: survey of an emerging domain," *Image and Vision Computing,* vol. 27, no. 12, 2009, pp. 1743-1759.

[2] J. K. Burgoon, N. Magnenat-Thalmann, M. Pantic, and A. Vinciarelli, *Social Signal Processing*. Cambridge, United Kingdom; New York, NY: Cambridge University Press, 2017.

[3] C. Geertz, "Thick description: toward an interpretive theory of culture," in *The Interpretation of Cultures*. New York: Basic books, 1973.

[4] R. Aylett and A. Paiva, "Computational modelling of culture and affect," *Emotion Review*, vol. 4, no. 3, 2012, pp. 253-263.

[5] P. Ekman, "Basic emotions," in *Handbook of Cognition and Emotion*, T. Dalgleish and M. Power, Eds. Chicester: John Wiley & Sons, 1999, pp. 45-60.

[6] P. Ekman, *Emotions Revealed: Recognizing Faces and Feelings to Improve Communication and Emotional Life*, 2nd Owl Books ed. New York: Owl Books, 2007.

[7] S. S. Tomkins and B. P. Karon, *Affect, Imagery, Consciousness*. New York: Springer Pub. Co., 1962.

[8] J. Panksepp and D. Watt, "What is basic about basic emotions? lasting lessons from affective neuroscience," *Emotion Review*, vol. 3, no. 4, 2011, pp. 387-396.

[9] A. R. Damasio, *Looking for Spinoza: Joy, Sorrow, and the Feeling Brain*, Orlando: Harcourt, 2003.

[10] Lutz, Catherine and Lila Abu-Lughod, eds. *Language and the Politics of Emotion*, Cambridge, UK: Cambridge University Press, 1990.

[11] R. I. Levy, "The emotions in comparative perspective," in *Approaches to Emotion*, K. R. Scherer and P. Ekman, Eds. New Jersey: Lawrence Erlbaum Associates, 1984, pp. 397-412.

[12] C. Lutz, *Unnatural Emotions: Everyday Sentiments on a Micronesian Atoll & their Challenge to Western T*heory, Chicago: University of Chicago Press, 1988.

[13] Z. Michelle, Rosaldo, *Knowledge and Passion: Ilongot Notions of Self and Social Life*, Cambridge, UK: Cambridge University Press, 1980.

[14] B. B. Schieffelin, *The Give and Take of Everyday Life: Language Socialization of Kaluli Children*, Cambridge, UK: Cambridge University Press, 1990.

[15] R. A. Shweder and R. A. LeVine, *Culture Theory: Essays on Mind, Self and Society*, Cambridge, UK: Cambridge University Press, 1984.

[16] M. Wetherell, *Affect and Emotion: A New Social Science Understanding*, Los Angeles; London: SAGE, 2012.

[17] R. J. Davidson, "Commentary—seven sins in the study of Emotion: correctives from affective neuroscience," in *Brain and Cognition*, vol. 52, no. 1, 2003, pp. 129-132.

[18] L. F. Barrett, *How Emotions are Made: The Secret Life of the Brain*, Boston: Houghton Mifflin Harcourt, 2017.

[19] D. White, "Affect: an introduction," *Cultural Anthropology*, vol. 32, no. 2, 2017, pp. 175-180.

[20] J. J. Lee, F. Sha, and C. Breazeal,"A Bayesian theory of mind approach to nonverbal communication," in 2019 14th ACM/IEEE International Conference on Human-Robot Interaction (HRI), 2019, pp. 487-496.

[21] K. Fischer, M. Jung, L. C. Jensen, and M. V. a. d. Wieschen, "Emotion expression in HRI—when and why," in 2019 14th ACM/IEEE International Conference on Human-Robot Interaction (HRI), 2019, pp. 29-38.

[22] M. F. Jung, "Affective grounding in human-robot interaction," in 2017 12th ACM/IEEE International Conference on Human-Robot Interaction (HRI), 2017, pp. 263-273.

[23] P. Ekman and W. V. Friesen, *Facial Action Coding System: Investigator's Guide*, Consulting Psychologists Press, 1978.

[24] T. Baltrušaitis, P. Robinson, and L. Morency, "OpenFace: an open source facial behavior analysis toolkit," in 2016 IEEE Winter Conference on Applications of Computer Vision (WACV), 2016, pp. 1-10.

[25] Mirzaei, M. S., Zhang, Q., Van der Struijk, S., and Nishida, T., “Language learning through conversation envisioning in virtual reality: a sociocultural approach, in P. Taalas, J. Jalkanen, L. Bradley & S. Thouësny eds., Future-Proof CALL: Language Learning as Exploration and Encounters – Short Papers from EUROCALL, 2018, pp. 207-213.

[26] P. Ekman, "Universals and cultural differnences in facial expressions of emotion," in Nebraska Symposium on Motivation, vol. 19, J. Cole, Ed. Lincoln: University of Nebraska Press, 1972.

[27] W. V. Friesen, "Cultural differences in facial expressions in a social situation: an experimental test of the concept of display rules," Dissertation Abstracts International, 33(8-B), 3976-3977, 1972.

[28] P. Ekman and W. V. Friesen, "The repertoire of nonverbal behavior: categories, origins, usage, and coding," *Semiotica*, vol. 1, 1969, pp. 49-98.

[29] J. Slaby, R. Mühlhoff, and P. Wüschner, "Affective arrangements," *Emotion Review*, 2017, pp. 1-10.

[30] R. Williams, *The Long Revolution*. London: Chatto & Windus, 1961.

[31] B. Massumi, *Parables for the Virtual: Movement, Affect, Sensation*, Durham: Duke University Press, 2002.

[32] B. d. Spinoza, *Ethics*, London: Penguin Books, 1996 [1677].

[33] M. Mori, K. F. MacDorman, and N. Kageki, "The uncanny valley: The original essay by Masahiro Mori," IEEE Spectrum, 2012.

[34] J. Robertson, *Robo Sapiens Japanicus: Robots, Gender, Family, and the Japanese Nation*, Oakland, California: University of California Press, 2018.

[35] K. Floyd and V. Manusov, "Biological and social signialing systems," in *Social Signal Processing*, J. K. Burgoon, N. Magnenat-Thalmann, M. Pantic, and A. Vinciarelli, Eds. Cambridge: Cambridge University Press, 2017, pp. 11-22.

[36] A. R. Damasio, *The Strange Order of Things: Life, Feeling, and the Making of the Cultures*, New York: Pantheon Books, 2018.

[37] B. Latour, *Science in Action: How to Follow Scientists and Engineers through Society*, Cambridge, Mass.: Harvard University Press, 1987.

[38] B. Latour, *Politics of Nature*, Cambridge, Mass.: Harvard University Press, 2004.

[39] P.-P. Verbeek, *What Things Do: Philosophical Reflections on Technology, Agency, and Design*, University Park, Pa.: Pennsylvania State University Press, 2005.

[40] H. Katsuno, "The robot's heart: tinkering with humanity and intimacy in robot-building," *Japanese Studies*, vol. 31, no. 1, 2011, pp. 93-109.

[41] C. Wagner, *Robotopia Nipponica: Recherchen zur Akzeptanz von Robotern in Japan* (Robotopia Nipponica: Research on Robot Acceptance in Japan). Marburg: Tectum Wissenschaftsverlag, 2013

[42] J. Wright, "Tactile care, mechanical hugs: Japanese caregivers and robotic lifting devices," *Asian Anthropology*, vol. 17, no. 1, 2018, pp. 24-39.

[43] C. Tsiourti, A. Weiss, K. Wac, and M. Vincze, "Multimodal integration of emotional signals from voice, body, and context: effects of (in)congruence on emotion recognition and attitudes towards robots," *International Journal of Social Robotics*, 2019.

[44] Groove X. Official website [Online]. Available: https://lovot.life/ [Accessed 1-May-2019].

[45] H. Gunes, M. Piccardi, and M. Pantic, "From the lab to the real world: affect recognition using multiple cues and modalities," in *Affective Computing: Focus on Emotion* Expression, Synthesis, and Recognition, J. Or, Ed. Vienna: IntechOpen, 2008, pp. 185-218.

[46] S. E. Kirksey and S. Helmreich, "The emergence of multispecies ethnography," *Cultural Anthropology*, vol. 25, no. 4, 2010, pp. 545-576.

[47] E. Kirksey, *The Multispecies Salon*, Durham: Duke University Press, 2014.

[48] D. J. Haraway, *Staying with the Trouble: Making Kin in the Chthulucene*, Duke University Press, 2016.

[49] P. Dumouchel, "Social emotions," in *Animating Expressive Characters for Social Interaction*, L. Cañamero and R. Aylett eds., Amsterdam: John Benjamins Publishing Company, 2008, pp. 1-19.