# Can We Have Cultural Robotics Without Emotions?

Candiotto, Laura; Mansouri, Masoumeh

*DOI:*
https://doi.org/10.3233/FAIA220625

*Document Version*
Peer reviewed version

*Citation for published version (Harvard):*
Candiotto, L & Mansouri, M 9, Can We Have Cultural Robotics Without Emotions? in *Can We Have Cultural Robotics Without Emotions?.* pp. 259. https://doi.org/10.3233/FAIA220625

# Can We Have Cultural Robotics Without Emotions?

Laura Candiotto[a,1] and Masoumeh Mansouri[b]
[a] *Centre for Ethics of the University of Pardubice*
[b] *School of Computer Science, University of Birmingham*

**Abstract.** As robots begin to operate in environments hitherto only occupied by humans, it has become common to introduce social and cultural thinking into robotics research. Starting from the conceptualisation of culture as an emergent property of participatory processes of sense-making, in this paper we argue that cultural robotics cannot be achieved without emotions. To defend this thesis, we will first justify our working concepts, then provide our argument along with clarificatory examples. We also point to limitations of the current state-of-the-art in realising an "emotional robot" required in cultural robotics, and finally, set out a roadmap for robotics research into emotion and culture.

**Keywords.** cultural robotics; emotions; social understanding; participatory sense-making; enaction; human-robot interactions.

## 1. Introduction

Ornelas et al (2022) define culture that counts for robotics as a phenomenon that emerges from complex interactions between body, environment, and other agents. According to this conceptualisation, what matters for cultural robotics is, at least initially, not a predetermined 'culture', but social dynamics and social learning that enable an agent to mutually participate in and co-create culture. Šabanović et al. (2014) also advocate a concept based on the co-construction of culture and scientific practice and technology design. Their core argument is that given the emergent nature of culture, it is not adequate for a robot to merely possess knowledge of the rules and norms of a certain culture to engage in culturally meaningful interactions. Instead, robots must become participants in the ongoing process of social appropriation and invention that re-creates our culture. Therefore, a robot must be designed with the capability to participate in the interactions that lead to the arising of cultural behaviour.

This conceptual framework challenges a univocal account of 'culture as nationality' applied in some current trends of cultural robotics (Lim et al. 2020, Wang et al. 2010, Eresha et al. 2013, Khaliq et al. 2018, Trovato et al. 2015), by focusing on the emergent nature of culture, especially with social interactions as focal point. Though aware this conceptualisation does not have wide acceptance, we do not argue for it here, but explore the implications, if it is valid, for the growing field of cultural robotics. In particular, we argue that without emotions we cannot have emergent,

[1] laura.candiotto@upce.cz, University of Pardubice, Centre for Ethics, Department of Philosophy and Religious Studies, Stavařov 97, 532 10 Pardubice, Czech Republic.

cutural behaviour arising from human-robot interactions. The reason is that emotions are a necessary component of social understanding.

By social understanding, we mean the agents' ability to comprehend each other in social interactions. Social understanding is a crucial feature of intersubjectivity, as the process through which agents share experiences, knowledge and expectations. There are many and various accounts of social understanding in different research fields and traditions of thought. For example, in the cognitivist tradition, theory theory (TT) and simulation theory (ST) are the most employed.[2] However, we posit that the enactive one, i.e. participatory sense-making (De Jaegher and Di Paolo 2017), is the one that best suits our working definition of culture because it conceptualises social understanding as emerging from the dynamics of interaction. Social interaction is made of embodied coordination patterns with specific dynamic signatures. These patterns are measurable and can be operationalised, but they also have an inextricable experiential dimension (Fuchs and De Jaegher 2009; Gallagher 2013). At the centre of the enactive approach is the way agents make sense of and with each other, co-determined by themselves, the other, and the interaction processes they engage in. All three of these elements — the (at least) two participants, and their interaction — are effective factors in intersubjective skills and lives.

Importantly, emotions play a crucial function in participatory sense-making (Candiotto 2019). Emotions like fear, wonder, shame, gratitude, and anger, just to name a few among the extremely rich and nuanced emotional landscape, are in-between agents by shaping interaction and conferring specific features on them. For instance, gratitude opens up to the interlocutor with a feeling of appreciation of their value. This feeling enables the agents to constitute a positive relationship that, in turn, can strengthen their affective bonds and disclose new joint action possibilities. On the contrary, fear creates a distance between the interlocutors and the related feelings of being in danger and mistrust can shape their relationships in terms of suspiciousness and avoidance. This means that the quality of their relationship is scaffolded by the interlocutors' emotional dynamics. Dynamics that are not fixed but continuously reshaped by the development of patterns of interaction. The interlocutors, in this way, make sense of themselves and their being together, in and through interactions.

Given robot cultural participation is the key to cultural robotics, we posit that cultural robotics cannot be achieved without emotions; and that the enactive account of social understanding is a crucial tool for fine-graining the role of emotions in the generation of culture, as an emergent property of human and robot interactions. In section 2, after a brief introduction of the current trends in cultural robotics, we focus on the role of emotions in cultural robotics. We criticize the reductionist stance of this research field while spelling out our enactive alternative in section 3. Our proposal is grounded on three core roles played by emotions in social understanding, namely orienting attention to salient features of the environment, informing and motivating empathing behaviours among the interlocutors, and evaluating the interactional experience. By the analysis of these three roles in the example of a robot who hands over a book, we claim that emotions are a necessary component of human-robot social

[2] TT contends that social understanding depends on a practice of mentalizing in which agents employ common sense about how mental states inform the behaviors of others. ST claims that agents can use their own mind to simulate the other person's mental states. For a discussion of the cognitive theories along with the exposition of alternative views, see Gallagher 2008.

understanding. In particular, we focus on the co-participation of humans and robots, as regulated by social affordances and successful interactions. Finally, in section 4, we discuss the benefits of employing this enactive account in cultural robotics, also stressing some crucial issues that deserve further exploration, such as embodiment and historicity.

## 2. Cultural Robotics and Emotions in Robotics Research

Lim et al. (2020) analysed 50 studies at the intersection of culture and social robotics. In all studies, culture was understood as national culture. Only a minority of research in cultural robotics differs from this dominant view. Šabanović (2014), Ornelas et al. (2022) and Winfield (2018) are among this minority that interprets culture as developing through interactions between social actors, including artificial actors. The common denominator of these studies is that culture is dynamic and situated in the environment; it is shaped via repeated interactions and co-developed.

The treatment of emotion in robotics is not so different from that of culture. The field of “emotional robotics” is overwhelmed by reductionist approaches in which the emotional experience is compartmentalised into universal and basic emotions such as fear, anger, disgust, sadness, happiness and surprise (Ekman, 1999) or similar discrete, often custom made, categories (Sreeja & Mahalakshmi, 2017). The discretisation of emotion is not the sole source of simplicity. Some models are defined through a continuous scale, with the most common being the circumplex model, mapping emotion on the two dimensions, valence and arousal (Posner et. al. 2005). Even by disregarding the existing critiques of the model in cognitive science, the usage of this model in robotics does not cover a wide range of overlapping emotional states, rather, to facilitate the algorithmic process of recognising "typical" emotions (e.g., joy, anger, sadness, and relaxation), for instance from a sequence of images via an artificial neural network, e.g., Tsujimoto et. al. (2016).

As mentioned, the existing limitations of realising emotion in robotics lie mostly in the reductionist and merely functionalist conceptualisation of the emotional experience. In a recent review article, Savery & Weinberg (2021) analysed 232 papers on the intersection of emotion and social robotics. Based on their extensive study, they conclude that “the majority of papers used emotion at most at a dyadic occurrence (one human and one robot), and often as a solitary experience functioning internally within the robot”, page 1037. Hence, the dynamics among the robot, environment and other agents of the environment are largely ignored. Also, in these studies, emotion is considered static, a snapshot of interaction devoid of any influence from experience. According to this survey, only 27% of papers had some form of history, the majority of which translate history to a single previous step. The absence of considering history in interaction has an implication in the realisation of culture in robotics, as it assumes that agents are interacting in a void. We will describe later, that this assumption contradicts the very properties of the enactive account of emotion that are essential for the implementation of cultural emergence.

## 3. Emotions for Cultural Robotics: An Enactive Account

The theory of cultural robotics this paper subscribes to is based on the emergence of culture through the interactions of several individual component parts of a process (Ornelas et al. 2022). This theory borrows foundational principles from embodied

cognition, ecological psychology, and dynamical systems research. The key ingredient of this theory is the formation of culture via repeated interactions, co-development, dynamic learning and coordination of all the parts involved. These include humans, robots and salient environmental features relevant to interactions. It follows that the emergent theory of cultural robotics based on co-participation involves (1) an environment model in which robots do not simply perceive the world as raw sensor data, but rather as social affordances, i.e. opportunities for interaction; (2) robots being able to learn and discover context-based cultural behaviour through successful and unsuccessful interactions.

Both points (1) and (2) cannot be implemented without emotion, as emotion serves the following functionality in participatory sense-making:

(a) emotions provide information about the participants and their relationships, by orienting attention to salient features of the environment (Brady 2018). For a robot looking for opportunities for interaction (1), there are infinite possibilities. Emotion affords the space of possibilities by discovering salient features of the environment for participation (Griffith and Scarantino 2009; Schargel and Prinz 2018). So, emotional affordances explain emotional motivation and behaviour. Importantly, this can take the shape of a joint commitment between humans and robots as well (Salice and Michael 2017). This is crucial to the robot's motivation to interact (1).

(b) emotions are constitutive of primary intersubjectivity and empathic behaviours of feeling bodies (Daly 2014; Fuchs 2016); hence without them, we cannot determine what successful and unsuccessful interactions, are as an intersubjective and embodied encounter between humans and robots (2).

(c) emotions are expressions of the agent's needs and contribute to sense-making (on affective sense-making, see Colombetti 2014). This means that the felt quality of emotion ("valence" in the technical jargon) expresses the agent's specific evaluation of the experience (Döring 2010), implicit reasons (Betzler 2007), attitude (Helm 2002), concerns (Fridja 1986), and meaning (Slaby & Stephan 2008); again, evaluation of the experience of interaction is what determines successful and unsuccessful interactions (2).

(1) and (2) are deeply interlaced and we argue that emotions play a significant role in both. In order to explain what we mean by successful and unsuccessful interactions and how they are shaped by the affective perception of opportunities for interaction, consider the example of a robot being tasked to hand over a book to a human. One might argue that a success criterion for this interaction is when the human receives the book robustly, i.e., the book is in the hand of the human. However, if the book is considered holy, for instance, due to its religious connotation, our expectation of handing over a book is clearly different. The successful interaction is the one book treated respectfully. Success/unsuccess criteria can vary depending on the environment, the object in question, the interlocutors' affective states, the common practice of handing over and many other factors. From an enactive point of view, both the interlocutors are co-participants in the carving out of their affective niches. In doing so, the evaluation of success is not only based on the result of the action (if the robot gives the book) but also the affectively charged existential needs that make them afford the situation in a certain way. Culture, or in our case religion as one of its instantiations, replies to existential needs, such as recognition, love, and social status. At the same time, culture is enacted in and through the affectively charged interactions that carve out and shape shared meanings. It follows that meanings are co-created by human and robot together in and through social understanding. So, the affective dimension of

social understanding constitutes the how of the interaction, i.e. its quality, in both its passive (response to existential needs) and active (sense-making) dimensions. The quality of interaction cannot be captured with a mere functionalist understanding of success. So, we posit that a focus on emotions in participatory sense-making can help in making less reductionist the work on cultural robotics that is so dependent upon successful interactions between humans and robots.

## 4. Discussion

In section 2, we briefly explained that the majority of research in robotics employs a reductionist approach in dealing with emotion. For instance, it is common in robotics experiments to ask the user of the system whether an interaction is successful or unsuccessful. This is often implemented by asking repeatedly a simple yes or no question over a different range of behaviours. Then based on scoring criteria, a behaviour with a higher score, for instance the one with the highest yes response, would be the outcome of the training process. This process, although explained in a simplified form, is the basis of the Reinforcement Learning algorithm, one of the most popular approaches in the use of AI methods in robotics. The range of behaviours in the 'handing over book' example mentioned earlier, can be spanned by the different ways a robot can grasp a book (contact points) and its distance with the interacting human. Clearly, in this instance, a reduction of interaction to a binary yes/no as well as lack of consideration of the context and the history of interaction falls short of exhibiting functionality (a), (b) and (c). In general, whether current methods in robotics and AI are computationally and epistemologically adequate for capturing emotion as participatory sense-making processes with the (a) to (c) functionalities featured above is a subject for future research. However, there is no reason to believe the impossibility of capturing emotion as such through the combination of various AI methods ranging from knowledge representation, automated reasoning and machine learning.

In particular, we posit that robots' embodiment is one of the main issues that needs to be faced by the implementation of an enactive model of participatory sense-making in cultural robotics. The reason is that for (b), primary intersubjectivity is always embodied. Embodiment does not only mean to have a physical body (*Körper*) but, in the case of embodied interactions, to experience them as a living body (*Leib*) with its needs, desires, concerns, and vulnerabilities.[3] And this points to (c), namely on how the felt quality of experience kindles processes of sense-making.

Regarding the realisation of emotions for robots from this perspective, Man and Damasio's (2019) proposal for the implementation of homeostasis (i.e., the regulation of body states) to soft robotics is an important contribution to the research, but the feasibility of getting from this proposal also the emotional embodiment required for developing robotic culture out of processes of participatory sense-making is something that needs to be carefully explored and assessed, both in its epistemological and applicative side.

A different answer to the issue of robots' embodiment in the emergence of robotic culture as participatory sense-making would be to argue that embodiment need not be necessarily bidirectional. It might be argued that it is enough that just one of the interlocutors, most likely the human, displays the level of embodiment we were

[3] For the phenomenological distinction between *Körper* and *Leib* and what a richer conceptualisation of embodiment entails, see Dreyfus 2009.

referring to above (Seibt 2017). But, in this case, we might wonder if the creation of culture would really emerge out of a *participatory* process of sense-making as we argued in this paper. The point here is to fine-grain what kind of participation is achievable between humans and robots. We believe that we do not need to assume that it should be the same as a participatory process of sense-making between humans. Moreover, participation should not be taken as a bidirectional of equals - it is possible to have participation between two different beings. In the case of the emergence of cultural robotics, culture is an emergent product of agents' interactions, e.g., among humans, robots, objects in the environment and the aggregation of learned coordination activities.

Another alternative would be to disjoint the physical body and the generation, recognition and expression of needs, desires, concerns and vulnerabilities, instead focusing on them from a disembodied perspective. Arguably, humans get to this important existential dimension of participatory sense-making through the body, but it is not said that this should be the case for robots as well. It might follow that in designing feeling robots *other* systems can provide access to these important features. The risk of this option is doing it in a reductionist fashion, but we do not want to exclude the possibility of doing it in a richer way.

The second point we want to stress is the role of historicity in the emergence of culture in the human-robot interaction. This is a core feature that the reductionist approach lacks in dealing with emotion in robotics. For the enactive account we employ here, agents are not interacting in a void. On the contrary, their embodiment is always situated in specific contexts that are already culturally laden.[4] Would that mean that the agents do not create a new culture through their interaction but simply absorb the contextual culture? If that would be the case, there would not be a real "emergence" of a culture between human and robots but the incorporation of pre-existing environmental features. We do not think that this is the case because participatory sense-making is a creation of new meanings in the interactions, and the interacting robot brings an interlocutor to the interaction for which not much historical knowledge exists. The point is that the interactions should not be understood in a narrow and reductionist way, but as dynamical processes between the agents and their shared environments. Specifically, this means considering the context as a factor that influences the interactional dynamics. This translates to focusing on the transformation of robotic culture as something that evolves and changes depending on the agents, their relationships and their shared worlds. It might be argued that the weight of the contextual factors is too loud and so the interlocutors would not have the power to create a new culture through their interaction. Although we think that the enactive model can point to the feasibility of this process because it focuses on the constant feedback loops between agents, their relationships and their environment, here we simply state that this important issue needs to be further investigated.

## 5. Conclusion

In this paper, we argue that viewing emotion as a necessary condition for participating in the co-creation of culture opens a new space of possibilities for the inclusion of emotion in robotics and artificial intelligence research. In this line of research,

[4] In particular the pragmatist take on enaction has stressed this point. See, for instance, Candiotto & Dreon 2021.

modelling of the context within which humans and robots operate as well as the social dynamics of their interaction, captured in their mutual experience, are crucial. These are the very properties that have been ignored in the current implementation of emotions in robotics (Savery & Weinberg, 2021). We claim that a non-reductionist account of emotions should have a primary role in this model as they constitute important features of social understanding. Our proposal is doing it with an enactive account of emotions in participatory sense-making. A critic might argue that a reductionist approach to emotions is better than nothing and easier to apply to robotics. But our point is that if we really want to develop cultural robotics, this would be detrimental because it would not take into account the emergent nature of culture. Moreover, we think that any suggestion about the possibility or impossibility of applying a non-reductionist theory of emotions to cultural robotics is baseless because, to the best of our knowledge, which is substantiated by a recent review of the field (Savery & Weinberg, 2021), no one has even tried it in the first place. So, the final upshot of our research is the necessity of grounding cultural robotics in a richer and more nuanced account of emotions in interactions.

## References


[1] Ornelas ML, Smith GB, Mansouri, M. Redefining Culture in Cultural Robotics. Journal of AI & SOCIETY. 2022.

[2] Brady M. Emotion: The basics. Abingdon: Routledge. 2018.

[3] Betzler M. Making Sense of Actions Expressing Emotions. dialectica. 2007; 61(3): 447-466.

[4] Candiotto L. Emotions in-between. The affective dimension of participatory sense-making, in The Value of Emotions for Knowledge, ed. L. Candiotto, 2019; pp. 235-260. London: Palgrave Macmillan. https://doi.org/10.1007/978-3-030-15667-1_11

[5] Candiotto L, Dreon R. Affective Scaffoldings as Habits. A Pragmatist Approach. 2021. Frontiers in Psychology. https://doi.org/10.3389/fpsyg.2021.629046

[6] Colombetti G. The Feeling Body. Affective Science Meets the Enactive Mind. 2014. Cambridge (MA): MIT Press.

[7] Colombetti G, Krueger J. Scaffoldings of the affective mind. Philosophical Psychology. 2014; 28: 1157–1176. doi: 10.1080/09515089.2014.976334

[8] Daly A. Primary Intersubjectivity: Empathy, Affective Reversibility, 'Self Affection' and the Primordial 'We'. Topoi, 2014; 33, pp. 227-241

[9] De Jaegher H, Di Paolo E. Participatory Sense-Making: An enactive approach to social cognition. Phenomenology and the Cognitive Sciences. 2007; 6(4): 485–507. doi: 10.1007/s11097-007-9076-9.

[10] Dreyfus HL. On the Internet. 2009. London-New York: Routledge.

[11] Döring SA. What a Difference Emotions Make. In Blackwell Companion to the Philosophy of Action, eds. T. O'Connor. C. Sandis. 2010; pp. 191-200. Oxford: Wiley-Blackwell.

[12] Fridja N. The Emotions. 1986; Cambridge: Cambridge University Press.

[13] Fuchs T. Intercorporeality and Interaffectivity. in Intercorporeality: Emerging Socialities in Interaction, eds. C. Meyer, J. Streek, S. Jordan; 2016; pp. 3-23. New York: Oxford University Press.

[14] Fuchs T, De Jaegher H. Enactive Intersubjectivity: Participatory sense-making and mutual incorporation. Phenomenology and the Cognitive Sciences. 2009; 8(4): 465–486. doi: 10.1007/s11097-009-9136-4.

[15] Gallagher S. Direct perception in the intersubjective context. Consciousness and Cognition. 2008: 535-543.

[16] Gallagher S. In the shadow of the transcendental: Social cognition in Merleau-Ponty and cognitive science. In Corporeity and affection: dedicated to maurice Merleau-Ponty, ed. K. Novotný, P. Rodrigo, J. Slatman, and S. Stoller. 2013: 149–158. Boston and Leiden: Brill.

[17] Griffiths P, Scarantino A. Emotions in the wild. the situated perspective on emotions, in The Cambridge Handbook of Situated Cognition, eds. P. Robbins and M. Aydede (New York, NY: Cambridge University Press). 2009: 437–453. doi: 10.1017/CBO9780511816826.023

[18] Helm B. Felt Evaluations. A Theory of Pleasure and Pain. American Philosophical Quarterly. 2002; 39(1): 13-30.

[19] Seibt J. Towards an ontology of simulated social interaction: varieties of the “As If” for robots and humans. InSociality and normativity for robots. 2017; pp. 11-39. Springer, Cham.
[20] Man K, Damasio A. Homeostasis and Soft Robotics in the Design of Feeling Machines. 2019; Nature Machine Intelligence https://doi.org/10.1038/s42256-019-0103-7
[21] Shargel D, Prinz J. An enactivist theory of emotional content, in The Ontology of Emotions, eds. H. Naar and F. Teroni (Cambridge: Cambridge University Press). 2018; 110–129. doi: 10.1017/9781316275221.007
[22] Slaby J, Stephan A. Affective intentionality and self-consciousness. Consciousness and Cognition. 2008; 17: 506-513.
[23] Trevarthen C. What Is It Like to Be a Person Who Knows Nothing? Defining the Active
[24] Intersubjective Mind of a New Born Human Being, Infant and Child Development. 2011; 20, pp. 119-135;
[25] Šabanović S, Bennett CC, Lee HR. Towards Culturally Robust Robots: A Critical Social Perspective on Robotics and Culture. Proceedings of HRI Workshop on Culture-aware robots. 2014.
[26] Salice A, Michael J. Joint Commitments and Group Identification in Human-Robot Interactions. Sociality and Normativity for Robots, eds. R. Hakli, J. Seibt J. 2017; pp. 179-199. Cham: Springer.
[27] Savery R, Weinberg G. Robots and emotion: a survey of trends, classifications, and forms of interaction. Advanced Robotics. 2021; 35(17), 1030-1042.
[28] Ekman P. Basic emotions. Handbook of cognition and emotion. 1999; 98(45-60), 16.
[29] Posner J, Russell JA, Peterson BS. The circumplex model of affect: An integrative approach to affective neuroscience, cognitive development, and psychopathology. Development and psychopathology. 2005; 17(3), 715-734.
[30] Sreeja P, Mahalakshmi G. Emotion models: a review. International Journal of Control Theory and Applications. 2017; 10, 651-657.
[31] Tsujimoto T, Takahashi Y, Takeuchi S, Maeda Y. RNN with Russell's circumplex model for emotion estimation and emotional gesture generation. In 2016 IEEE Congress on Evolutionary Computation (CEC) (pp. 1427-1431). 2016, July. IEEE.
[32] Lim V, Rooksby M, Cross ES. Social robots on a global stage: establishing a role for culture during human–robot interaction. International Journal of Social Robotics, 2020; 1-27.
[33] Wang L, Rau PLP, Evers V, Robinson BK, Hinds P. When in Rome: the role of culture & context in adherence to robot recommendations, in Human-Robot Interaction (HRI), ACM/IEEE Int. Conf, 2010; pp. 359–366.
[34] Eresha G, Häring M, Endrass B, André E, Obaid M. Investigating the influence of culture on proxemic behaviours for humanoid robots, in Robot and Human Interactive Communication (RO-MAN), IEEE Int. Workshop, 2013; pp. 430–435.
[35] Khaliq AA, Köckemann U, Pecora F, Saffiotti A, Bruno B, Recchiuto CT, ... & Chong, N. Y. Culturally aware planning and execution of robot actions. In 2018 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS); 2018; pp. 326-332. IEEE.
[36] Winfield, AFT. Experiments in artificial theory of mind: From safety to story-telling. Frontiers Robotics AI, 2018; 5(JUN).