Elsevier Editorial System(tm) for Physics of Life Reviews

Manuscript Draft

Manuscript Number: PLREV-D-19-00019

Title: Implicit Perception Simplicity and Explicit Perception Complexity in Sensorimotor Comunication. Comment on "The body talks: Sensorimotor communication and its brain and kinematic signatures" by G. Pezzulo et al.

Article Type: Comment

Keywords: Active Perception; Social Perception; Theory of Mind; Perspective Taking

Corresponding Author: Dr. Dimitri Ognibene, Ph.D.

Corresponding Author's Institution: University of Essex

First Author: Dimitri Ognibene, Ph.D.

Order of Authors: Dimitri Ognibene, Ph.D.; Giuseppe Giglia, PhD; Letizia Marchegiani, PhD; David Rudrauf, PhD

Implicit Perception Simplicity and Explicit Perception Complexity in Sensorimotor Comunication. Comment on "The body talks: Sensorimotor communication and its brain and kinematic signatures" by G. Pezzulo et al.

Dimitri Ognibene ,1 School of Computer Science and Electronic Engineering, University of Essex, UK and 2 ETIC, Univeristat Pompeu Fabra, España (dimitri.ognibene@essex.ac.uk)

Giuseppe Giglia, 1. University of Palermo, Department of Biomedicine, Neuroscience and Advanced Diagnostics, Palermo, Italy 2. Department of Cognitive Neuroscience, Maastricht University, Maastricht, Netherlands (giuseppe.giglia@unipa.it)

Letizia Marchegiani, Department of Electronic Systems, Aalborg University, Denmark (lm@es.aau.dk)

David Rudrauf, School of Psychology, University of Geneva, Switzerland (David.Rudrauf@unige.ch)

The article by Pezzullo et al [1] has comprehensively addressed several aspects of Sensory Motor Communication (SMC) related to the complexity and the intentionality of processes on the information "sender" side. However, it pays less attention to the role of the same factors on the "receiver" side. It is known that interpretation of the detailed kinematics underlying the receiver side of sensorimotor communication is a non-trivial task. It requires the integration of detailed sensory information from multiple dynamic sources, such as the environmental context and the affordance it may entail, the relevant body parts and involved effectors [2]. Furthermore, those sources may not be all simultaneously accessible to sensory systems, e.g. being out of the field of view, warranting the adoption of active perception strategies [3-5]

Surely other processes contribute on the receiver side to realize SMC. Mechanisms of social perspective taking, i.e. the ability to mutually understand, even in a probabilistic sense, others interest, bound and facilitate the receiver task in SMC by helping forming predictions and relevance weighting about the different input sources and possible targets of interest. Thus, building expectations about specific contexts in SMC can be supported by higher-order forward models of spatial intentionality, playing as a common frame of interpretation assumed by the interacting agents. The detailed observation of the actual kinematics as well as other information sources, which are part of the more general social context, surely can be used to induce a revision of these expectations. But more abstract priors about the relevant information to extract, can reduce sensory processing load and facilitate a complex task. At minima, simple indications of overt spatial attention and orientation, e.g. yielding 3D look at information, combined with emotion expression, e.g. indicating whether to approach or avoid a location, can greatly contribute to reduce the complexity, in terms of sensory information integration, of social interactions and coordination tasks. It has been proved, for instance, that, while the orientation of covert spatial attention in the auditory domain is highly affected by the presence of a task [6], it can be facilitated by the emotional valence of auditory stimuli, even when task-irrelevant [7, 8]. In the same way, the cognitive computation demands evoked by social interactions can revise not only attention distribution, but the neural correlates of processing speech as well [9].

Thus regarding the distinction between true SMC and information transfer, according to pragmatics (i.e. the ability to use language, gestures and body movements to convey communicative meaning in a context [10]), it is worth noting that at least an understanding of partner's mental state (i.e. a theory of mind) should be present in both the receiver and the sender [11]. Experiments with human-robot interactions in which tasks requiring SMC are performed could add insight to the topic while observing to which extents humans use SMC to communicate with humanoid vs non biologically inspired industrial robotic arms (as an example see Oztop et al, 2005 [12]).

Thus we propose that the observer may often adopt simpler strategies during interaction and deal with the interpretation of the detailed kinematics only when deemed necessary. In such cases, information sampling becomes a very active process where the receiver and the sender are strongly involved.

Back to the initial explanatory experimental condition of moving a table jointly, we can expect two different types of observer behaviours. In one case (*serendipitous or implicit*) the receiver will decide only on the basis of the current state of the table without consideration for the intentions of the "sender", even though those will still affect the "receiver" actions by changing the state of the table. In the other condition (*engaged or explicit*), the "receiver" will try to collect the signaling information provided by the 'sender', integrate it over time, and infer the intentions of the sender. This can be done in a hierarchical manner starting from parsimonious information sampling.

The difference can be seen more explicitly in a simplistic example where the receiver is pushing the table towards the right while the sender is pushing it in the opposite direction. In the first, simple *serendipitous* case, we can expect that the receiver policy will either be independent from the force applied by the sender (i.e. the receiver will simply continue pushing), or dependent on it (i.e. the receiver will follow the lead of the sender). In the second, complex *engaged* case, the opposing reaction of the collaborator may initially raise surprise and uncertainty in the receiver, which, consequently, may wait for additional information, as the received signal mismatches with his/her expectations about the sender intentions, leading to the problem of re-estimating them. This may be achieved either by waiting or by actively probing different alternatives. This will also raise the receiver's attention towards the sender's efforts to improve his/her understanding. This example highlights the importance not only of signaling but of receiving the sender state, and alter it if necessary. A number of works in robotics tackling similar tasks [13–15] employ analogous collaboration strategies, each presenting a different degree of complexity.

In SMC based collaborations, both these observer's behaviours are likely to be present. The switch between the two is foreseeably regulated by the observed interaction progress, combined with, and related to, the observer's expectations. For instance, expecting a specific partner behaviour with high certainty will more likely lead to an observer behaviour of the first (*serendipitous*) type compared to a condition where the observer does not have clear expectations about the partner behaviour, which will more likely induce an observer behaviour of the second (*engaged*) type. When a large violation of the expectations is observed, a switch from the simple behaviour to the complex one is likely to occur.

In most situations, these roles have to be interpreted as mutually coupled: each agent is both a receiver and a sender. Interestingly, the difficulty of the SMC observation aspect as well as the different levels of effort (*serendipitous*, *engaged*) that the observer may engage in an SMC collaboration are likely to induce in the "sender" not only an increase of signaling efforts but also active check of the observer state to better adapt its signaling strategies. In fact, the sender might need to monitor the actual state of the receiver, as to produce a non-obvious collaborative maneuver he/she needs an engaged partner.

Overall, the complexity of the observation task is likely to induce in both the observer and the receiver complex active observation strategies to identify the partner intention and state. Ultimately, the involved epistemic actions [16,17] may be the most cognitively demanding components of an SMC process and will be so tightly coupled with the other parts of the interactions that we may consider SMC as a form of joint active inference. While several models capable of generating social epistemic actions to minimize surprise and uncertainty on partners intentions [2–5] have been proposed, these are quite computationally demanding and rigid, while the richness, flexibility and swiftness of the observed interactions point to the necessity of further studies to better understand the mechanism of active social perception. Integrating higher-order forward spatial models of intentionality based on perspective taking, could help resolve a great deal of complexity, by focusing information sampling on the most relevant cues. In the context of SMC, it would be interesting to study how the observer behaviour type (*serendipitous vs engaged)* affects the sender behaviours, both in terms of signaling and of monitoring of the observer. Using robotic observers would allow replicability, as well as the possibility of altering the sender expectations on the observer receptive skills.

References

[1] G. Pezzulo, F. Donnarumma, H. Dindo, A. D'Ausilio, I. Konvalinka, C. Castelfranchi, The body talks: Sensorimotor communication and its brain and kinematic signatures, Physics of Life Reviews (2018) 10.1016/j.plrev.2018.06.014

[2] Ognibene D, Chinellato E, Sarabia M, Demiris Y. Contextual action recognition and target localization with an active

allocation of attention on a humanoid robot. Bioinspir Biomim 2013;8:035002.
[3] Ognibene D, Demiris Y. Towards Active Event Recognition. IJCAI '13 Proceedings of the Twenty-Third International joint conference on Artificial Intelligence 2013: 2495-2501
[4] Lee K, Ognibene D, Chang HJ, Kim T-K, Demiris Y. STARE: Spatio-Temporal Attention Relocation for Multiple Structured Activities Detection. IEEE Trans Image Process 2015;24:5916–27.
[5] Donnarumma F, Costantini M, Ambrosini E, Friston K, Pezzulo G. Action perception as hypothesis testing. Cortex 2017;89:45–60.
[6] Marchegiani, L., Karadogan, S. G., Andersen, T., Larsen, J., & Hansen, L. K. (2011, December). The role of top-down attention in the cocktail party: Revisiting cherry's experiment after sixty years. In *Machine Learning and Applications and Workshops (ICMLA), 2011 10th International Conference on* (Vol. 1, pp. 183-188). IEEE.
[7] Asutay, Erkin, and Daniel Västfjäll. "Negative emotion provides cues for orienting auditory spatial attention." *Frontiers in psychology* 6 (2015): 618.
[8] Marchegiani, L., & Fafoutis, X. (2015). On cross-language consonant identification in second language noise. *The Journal of the Acoustical Society of America*, *138*(4), 2206-2209.
[9] Rice, K., & Redcay, E. (2016). Interaction matters: A perceived social partner alters the neural processing of human speech. *NeuroImage*, *129*, 480-488.
[10] Bosco FM, Tirassa M, Gabbatore I. Why Pragmatics and Theory of Mind Do Not (Completely) Overlap. Front Psychol 2018;9. doi:10.3389/fpsyg.2018.01453.
[11] Spanoudis G. Theory of mind and specific language impairment in school-age children. J Commun Disord 2016;61:83–96.
[12] Oztop E, Franklin DW, Chaminade T, Cheng G. HUMAN–HUMANOID INTERACTION: IS A HUMANOID ROBOT PERCEIVED AS A HUMAN? Int J Humanoid Rob 2005;02:537–59.
[13] Mörtl A, Lawitzky M, Kucukyilmaz A, Sezgin M, Basdogan C, Hirche S. The role of roles: Physical cooperation between humans and robots. Int J Rob Res 2012;31:1656–74.
[14] Madan CE, Kucukyilmaz A, Sezgin TM, Basdogan C. Recognition of haptic interaction patterns in dyadic joint object manipulation. IEEE Trans Haptics 2015;8:54–66.
[15] Oguz SO, Ozgur Oguz S, Kucukyilmaz A, Sezgin TM, Basdogan C. Haptic negotiation and role exchange for collaboration in virtual environments. 2010 IEEE Haptics Symposium, 2010. doi:10.1109/haptic.2010.5444628.
[16] Friston K, Rigoli F, Ognibene D, Mathys C, Fitzgerald T, Pezzulo G. Active inference and epistemic value. Cogn Neurosci 2015;6:187–214.
[17] Ognibene D, Baldassare G. Ecological Active Vision: Four Bioinspired Principles to Integrate Bottom–Up and Adaptive Top–Down Attention Tested With a Simple Camera-Arm Robot. IEEE Trans Auton Ment Dev 2015;7:3–25.