
SARAH PAWLETT JACKSON

St Mellitus College and University of London

sarah.pawlett-jackson@stmellitus.ac.uk

THREE BODIES: PROBLEMS FOR VIDEO-CONFERENCING¹

abstract

In this paper I examine a specific way that video-conferencing modifies structures of intersubjective awareness and interaction. I focus on multi-person interactions (involving more than two people) via video-call. By unpacking some of the key features of multi-person intersubjectivity in cases of embodied co-presence, I will show where and how certain social affordances are strained or lost when multi-person interactions are transferred to the screen.

keywords

intersubjectivity, phenomenology, second-personal awareness, video-chat

¹ The author wishes to thank two anonymous reviewers for their helpful comments and additional suggested reading which have helped to improve the paper.

1. Introduction Digital platforms shape, scaffold and modify our identities and experiences in many ways. Of these, modification of our *intersubjective interactions* with others are of real significance. With the increasing ubiquity of digital forms of interaction, particularly in the light of the COVID-19 pandemic and new norms of physical distancing, it is important to reflect on some of the specific ways that digital interfaces shape intersubjective and collective experiences. The meanings and possible meanings of these modifications need to be analysed and understood. In this paper my interest is specifically in the effects of video conferencing on multi-person forms of intersubjectivity. Multi-person intersubjectivity refers to personal interaction between more than two people. These are interactions that involve more than a single self-other dyad but are of a small enough size that interpersonal awareness between all parties can be maintained, namely interactions in small groups. Video-conferencing platforms, such as Skype or Zoom, make provision for multi-person conversations of this kind, with multiple people present and interacting on the same call. My objective is to draw attention to some specific ways that video-conferencing modifies forms of interpersonal awareness in these multi-person cases. I argue that social affordances that are ordinarily available in multi-person cases are not available in the same way via video.

This phenomenological interrogation is particularly important as multi-person intersubjectivity has been underexplored as a phenomenon in its own right. It is often implicitly (or explicitly) assumed that multi-person intersubjectivity can be understood as though it were the mere addition of various self-other dyads. Similarly, existing literature which considers the nature of modification to intersubjective communication via video has tended to consider factors that apply across dyadic and multi-person interactions, not looking specifically at the phenomenological modifications to small group video interactions *qua* small group.

I start by laying out some of the ‘social affordances’¹ that are ordinarily available to us in small group contexts of embodied “co-presence” (Gilbert 2014, p. 325) which involve more than two people. I will highlight how these affordances are grounded in embodied and enactive awareness. I will then look at how the structure of these forms of intersubjective awareness is typically modified when interacting via video calls. I will conclude that the inability to engage

¹ The language of affordances originally comes from Gibson (1986/2015). I will explain this phenomenon in more detail below.

with interactions-between-others is the most significant curtailment to organic small group interaction in online contexts.

The philosophy of intersubjectivity has tended to focus on forms of interaction within a self-other dyad. Typically, this has focused on either the first-person plural phenomena of joint attention and joint action (Hobson 1989; 2008; Eilan *et al* 2005; Seemann 2012; Zahavi 2014; Reddy 2015), or on second-person singular forms of direct “mutual recognition,” (Gilbert 2014, pp. 329-331), “mutual attention” (Reddy 2005. cf. Darwall 2006; De Jeagher & Di Paolo 2007; Reddy 2008; Stawarska 2009; Eilan 2014, Zahavi 2014: 2015), or shared emotion (Reddy 2008). Elsewhere, I have argued that cases of multi-person intersubjectivity, involving more than two people, should not be understood as mere linear extensions or additions of these dyadic interactions (Pawlett Jackson 2018; 2019). Rather, the complexities that emerge when interactions occur in the context of *other* interactions need to be taken into account. For example, I have argued that the *second-person plural* should be understood as a distinct form of interpersonal awareness in its own right, rather than a linear extension of the second-person singular. The second-person plural is not simply a grammatical category, but a specific form of intersubjective awareness in which I perceive and engage with multiple others *as a relational unit*.

For example, I might say to two friends of mine, a couple who have hosted me for the evening – ‘thank you so much for your hospitality’. All three of us, it seems, are aware of one another as subjects, and not only this, but this awareness is reciprocal in all directions. It is transparent to all that we are all aware of one another and all aware of one another’s awareness. This is not the whole story, however, because there is also an *asymmetry* in our awareness of one another that we are all aware of: I am also aware of them *as a pair* in a particular kind of way.

In this situation I am aware of both of my friends as distinct – they do not blur into a generic mass in my awareness, as might happen at a football match. Neither do they appear to me as isolated from one another. I experience them as individuals, but I *also* experience them *as a couple*. This is of course framed and clarified by an understanding of their existing historical relationship over time, but this is brought into and includes my awareness of them *as in relation to one another* in this moment in time and space, around the breakfast table. Minimally, and importantly for understanding the intersubjective structures in play, I am aware of them as peripherally but reciprocally aware of one another in their joint attention (as a ‘we’) *directed towards me*. I perceive a “conjoint availability” (Fivaz-Depeursinge & Corboz-Warnery 1999, p. 6) in the two of them which makes it possible and meaningful to address them as a pair. My awareness of their reciprocal awareness of one another is required if my address is to be understood by all parties as a “double address” (Fivaz-Depeursinge & Corboz-Warnery 1999, p. 114). This is precisely what seems to be in play when it is clear to all parties that I am addressing the two of them *together* with thanks. In such cases, I am able to address a pair or a small group as ‘you’ – that is *you plural*. All parties can recognise and understand what is happening in this kind of interaction – the second-person plural awareness is – in a phrase used by both Naomi Eilan (2005) and Margaret Gilbert (2014) – “out in the open” between us, “manifest to [all] participants” (Eilan 2005, p. 1).

This kind of intersubjective awareness, understood from the ‘other side’ of the interaction, as it were, reveals that it is possible to *be addressed* or to *address* as a relational unit. As part of a first-person plural – a “plural subject” to use Gilbert’s (2014) term² – we can jointly attend

2. Multi-person intersubjectivity

2 My use of this term does not mean that I am necessarily committed to every aspect of Gilbert’s account. I simply use it to refer to groups structured by joint attention.

to and address a third person together. A key principle here is that intersubjectivity is not just about interacting with others, but also about interacting with the *relationships between others* as well. This is why there are many forms of multi-person interaction which cannot be sufficiently understood as the mere addition of multiple second-personal dyads.

As I will consider further in the next section, our ability to recognise and participate in complex networks of interactive and asymmetrical jointness and mutuality requires sophisticated co-ordination. This does not mean, however, that this complexity of parts is salient in our lived experience. Indeed, we find that quite the opposite is true: we typically experience this reciprocity of awareness without needing to put in much attentional effort. This shows up in the way we perform complex coordination of our activity and communication with others without needing to think much about it. As Ackerman & Bargh (2010) note, “for many complex tasks...the ability to effectively coordinate with others requires intensive training. However, social coordination also occurs automatically, nonconsciously and effortlessly throughout our daily encounters with other people” (p. 335). Our ability to co-ordinate attention and action with others is grounded in pre-reflective forms of embodied awareness. I turn now to elaborate the role of ‘intercorporeality’ in multi-person interactions, before turning to look at video-conferencing in the light of this understanding.

3. Embodied and enactive

In recent years, phenomenologists and cognitive scientists have converged in understanding the human subject as inescapably ‘4E’ – *embodied, embedded, enactive and extended* (Varela *et al* 1991; Newen *et al* 2018). For the purposes of this paper I focus on subjectivity as embodied and enactive. As enactive beings the world presents itself to us not just as data for comprehension but as possibilities for action – or ‘affordances’.³ An understanding of attention as embodied and enactive is significant in our thinking about intersubjective interactions. Here, the explosion of recent work on enactive social perception underlines ideas originally laid down by Merleau-Ponty that ‘intercorporeality’ is the foundation of intersubjectivity (Morganti *et al* 2008; De Jaegher & Di Paolo 2007; Gallagher 2014; 2017; De Jaegher *et al* 2017; Fuchs 2017). This fundamental level of connection to others is, Merleau-Ponty (1945/2012) says, “achieved through the reciprocity between my intentions and the other person’s gestures, and between my gestures and the intentions which can be read in the other person’s behaviour” (pp. 190-192). In this vein, Francesca Morganti (2008) notes that traditional models “conceive of communicative interaction between two people as a process that takes place in a set of mental operations that end up being expressed (externalised) in speech, gesture and action” (p. 8). Embodied and enactive approaches, on the other hand, understand gestures, postures, bodily orientation and facial expressions to be substantive parts of intersubjective attention and intersubjective affectivity itself.

Ordinarily, I am aware of some components or combination of the other(s)’ posture, gesture or expressions. I experience the other’s attention directed at me in her shift in bodily orientation, for example. That (and how) I am held in the other’s attention might be clear to me in the openness of her posture, the hesitancy of her gestures or the thoughtfulness of her facial expression, as well as – and sometimes in spite of – what is communicated verbally.

This is the basis of the pre-reflexive coordination of bodies in space. Thomas Fuchs (2017) identifies this in the “inter-bodily resonance” that provides the basis of intersubjectivity, highlighting that when “mutual incorporation” is in play, “each lived body with its sensorimotor body schema reaches out, as it were, to be extended by the other” (p. 8). Fuchs highlights that this includes the “inter-affective” resonance between bodies. This can be seen

³ See fn1.

most explicitly in high-level coordination in joint kinaesthetic activities like sport or dance, but it is also found in the micro-exchanges of glances, movements and posture.

To communicate mutuality of awareness verbally is already 'too late' for ordinary forms of second-personal interaction, as it were. This communicative mutual awareness is the condition of the possibility of addressing second-personally at all. The embodied and extra-linguistic components of second-person mutuality make up what Stawarska (2009) calls the "deep grammar" or "protogrammar" (pp. 72; 130) of second-personal interactions.

To return to multi-person interactions specifically, we find that extra-verbal, embodied and enactive elements of recognition and communication are crucial in creating the conditions for the kinds of complex interlocking intersubjective awareness identified above. The ability to form a 'joint address' or to receive a 'double address' requires that I am able to coordinate my own actions, attentiveness and responses to and with multiple others at the same time. Significantly, I am able to recognise intersubjective interaction *between others in their* bodily coordination with one another. I can 'see' inter-affective resonance between others – I am able to recognise that the conversation between a couple is becoming tense, for example, or I can see that they are at ease between themselves and are inviting me to join their conversation. Where multiple perspectives within and between subjects and plural subjects are co-ordinated in the right way, seamless bodily and conversational coordination is possible between myself and multiple other subjects.

Unpacking this multiplicity, we can note that a single action, such as a gesture, can be a communication to more than one person at once. Not only this, but the body-language with which I communicate (and by which others communicate to me and to other others) is not limited to one kind of action. Gesture, posture, tone and facial expressions all play a role, as do multiple sensory modalities. Elizabeth Fivaz-Depeursinge and Antoinette Corboz-Warnerly (1999) identify the orientation and coordination of *pelves, torsos, gazes* and *expressions* as relevant in multi-person communication. They identify that these different forms of bodily orientation and expression can be coordinated in multiple ways, producing different kinds of three-way interaction (p. 57).⁴ I can also employ different bodily actions in a coordinated way to target more than one other person at once. When faced with multiple others I might, at any one time, be listening to one while looking directly at the other, for example, or looking at one while reaching out to the other with an inclusive gesture.

We can note specifically that I am able to coordinate attention *with* someone *to* another by coordinating different modes and actions. As *we face you*, I may gesture to my co-attendeé, shifting my posture to implicate her in my attentiveness – whilst looking at the one we are attending to together. There are some cases in which trying to coordinate attention to more than one other person with more than one modality could be confusing or distracting, causing a split in my or our attention. In other cases, however, these multiple sensory inputs may be integrated into a smooth whole – as part of an interaction that all three of us are attuned to as participants. As someone receiving a joint address, I can 'read' the intersubjective awareness *between the others* in the plural subject which I face. The reason I can do this is because it is enacted between them bodily. Their joint attention is not locked away from me as a private affair between their two (or more) disembodied minds but is out in the open for me to perceive and engage. Indeed, their attention to and with one another presents itself to me as a particular kind of social affordance for my second-personal engagement with them (together).

That the others' interlocking activity presents itself to me as an affordance or an invitation for engagement is important for understanding how various forms of asymmetrical multi-person interaction can emerge. It is not just a third-person plural observation of others' behaviour

⁴ This study looks at mother-father-infant triads as they play together.

that is in view here. I can respond and participate, not just recognise. I watch two people both writing at the same table in the café, for example. I would like to use a seat on the table and want to ask if I can take it. Are they there together? Should I address them separately or jointly? It is hard to tell at first, but as I take a moment to watch them, I recognise that they are there together by the way they are both holding their bodies, their orientation and their posture at the table. (Or indeed, I might recognise that they are there independently, functionally oblivious to one another.) This transforms how I address them with my request. Having laid out some key features of multi-person intersubjectivity in typical cases of co-presence, I turn now to look at multi-person interactions via video-conferencing. In considering the phenomenology of interactions via video I hope to illuminate specific social affordances that are strained or even lost in these contexts.

4. Others apart: multiplicity via video

In this section I focus on video-conferencing as a specific way that our intersubjective landscape can be mediated and therefore modified by digital technology. Literature already exists looking at a number of different ways that video technology modifies the phenomenology of intersubjective communication along axes such as empathy, affectivity, privacy, intimacy or objectification (Fuchs 2014; Ferencz-Flatz 2019; Krueger & Osler 2019; Osler Forthcoming). However, the significance of modifications to interpersonal structures of awareness in specifically multi-person contexts remains underexplored. My focus here is to note that there is a dyadic paradigm built into the architecture of the video call, and that certain key characteristics of organic multi-person dialogue are therefore impeded. When contrasting video calls with conversation that takes place in embodied co-presence, we can note that spatial and physical proximity is modified. Participants on a video call are temporally co-present to each other, but not spatially so.⁵ It is worth noting that there are other forms of technologically mediated multi-person interaction which are modified in different ways due to their different form. Group text messaging, for example, can involve multiple others who are dispersed not only in space but also in time. Some of the same principles of modification may apply to both video-conferencing and to group messaging, but there will be further issues to explore on the phenomenon of group messaging which I do not have space to look at here. I focus here on the way that spatiality is modified on video calls, and how this constrains embodied and enactive social affordances that we take for granted in co-located multi-person contexts.

In the words of product designer and software engineer John Palmer (2020), “video calls are faces inside static rectangles”. Not only this, but these static rectangles are stacked around each other on my screen in a way that does not and cannot take into account any relation between each of these others. The spatiality of the video call is a literal flattening out of those that I am speaking to – no one is nearer or further away from me or from anyone else. No-one can orient themselves to gesture at any other particular person, as we are not arranged together in shared space. To quote Palmer again:

When in real life are you ever looking at a grid of faces? Never...In real life, you look at people around a dinner table, or around a circle at...a meeting, or at someone beside you as you go for a walk (Palmer 2019).

5 As an aside, we can note that the prevalent phenomenon of glitches in internet connection often and typically modifies the temporal structure of a conversation over video as well. This is an interesting and important issue to explore for understanding modifications to interpersonal exchanges via video more generally. I will not focus on this issue here, staying with issues facing modified multi-person exchanges specifically. See Osler (Forthcoming) Section 5.2, for a discussion of the temporality of video-calls.

As a result of this modified spatiality, it is not possible to shift one's posture or orient one's body towards a particular person or sub-group of others within the group as a whole. Where gestures are made, they can only be made as a way of communicating to all interlocutors in the same way. If I try to orient towards or gesture to a particular person on the call, it is not transparent to this person, or anyone else on the call, that it is *this* particular person who I am gesturing to, save if I use language explicitly to single them out. I cannot catch a particular person's eye, not only because video calling makes this inherently difficult, even in dyadic cases, but because it is not possible for others to tell who it is on the call that I am trying to directly engage by my gestures or expression. This means that I cannot form the awareness of jointness or mutuality with particular people or sub-groups within the group – or at least not in the more organic way I would be able to do in a fully embodied context.

Likewise, because each of the others is cut off from one another in the same way, there is no bodily interaction between others in the group for me to 'read' and engage with. Whereas ordinarily I can 'pick up' the intercorporeal, inter-affective interactions between others in the ways that they orient and gesture towards one another, or share looks and facial expressions – none of this is available on a video call. I therefore cannot interact with others as a relational unit (from which I am distinct) in the same pre-reflective way. This particular form of social affordance is shut down when interacting via video.

While all participants share a joint awareness that everyone else is on the call, this is in a much more symmetrical or 'flattened' way. While in 'real life' there are various asymmetries to our awareness of each of the others – asymmetries that other others can recognise and then in turn interact with – this is not the case on a call. The 'conjoint availability' presented to me by others is far more limited. Each 'other' is atomised from the other others. On video calls we therefore experience a first-person plural experience between all of us participating, but this 'jointness' is the same in all directions, so to speak.

This is not to say that there is no scope for interacting with others as a relational unit whatsoever. We can still pick up tone of voice and detect conversational patterns between others. Particularly in established group relationships, with habituated forms of interaction in these relationships, it may be easier for us to pick up when two friends are sharing an in-joke together on a call, for example. Opportunities to recognise inter-affectivity between others are not completely shut down then, but they are significantly muffled. Video-calls need not involve a complete or absolute 'flattening' of asymmetrical dynamics. The point remains, however that the (non)spatial (non)arrangement of subjects means that I cannot organically form interactional sub-units within the wider group as I would ordinarily be able to do through my embodied and enactive co-ordination. It is not possible to align bodies with another person to speak as a 'we' to a third 'you', for example. Ordinarily in a group of three or four, conversation may naturally shift in its form from each participant speaking as an individual subject to parts of the conversation where two are addressed and jointly reply to the rest of the group as a plural subject. This kind of shift in intersubjective awareness, rooted as it is in embodied and enactive awareness, is much harder to access via video. As the key extra-verbal ways that we indicate jointness or mutuality are strained or lost, so forms of joint or mutual awareness can only be communicated explicitly and verbally. While this means that forms of joint address are still possible, ("X and Y, can I get your opinion on Z based on your experience of this together last year?"), this serves to make salient that the embodied and enactive scaffolding that usually grounds them is missing. More work is needed to make explicit what would otherwise be more immediately evident.

5. Conclusion In the above I have argued that there are specific forms of multi-person interaction which we ordinarily experience and participate in. These forms of interaction should not be understood reductively, as the mere addition of adjacent or overlapping dyads. This non-reductive understanding of multi-person interaction allows us to see that there are specific social affordances which video-conferencing curtails. Having laid out these specific structural inhibitions, it is hopefully clear that the claim that video-conferencing does not sufficiently enable forms of multi-person intersubjective awareness is not a vague or misplaced romanticism about ‘really being together’. Rather, this is an observation about the loss of multiplicity which is embedded in embodied co-presence. Going forward, this observation presents an opportunity for interface designers: what might it look like to build video-conferencing which builds shared spatiality and the possibility of asymmetrical intersubjective reciprocity into the technology itself? Future design work which takes seriously the non-reducibility of many forms of multi-person intersubjectivity may open up social affordances not yet possible via video technology.

REFERENCES

- Ackerman, J.M. & Bargh, J.A. (2010) Two to Tango: Automatic Social Coordination and the Role of Felt Effort. In Bruya, B. (ed.) *Effortless Attention* (pp. 335-371). Cambridge, Mass: The MIT Press; Darwall, S. (2006) *The Second-Person Standpoint*. Cambridge, Mass; London: Harvard University Press;
- De Jaegher, H. & Di Paolo, E. (2007) Participatory sense-making: An enactive approach to social cognition, *Phenomenology and the Cognitive Sciences*, 6(4), 485-507. doi: 10.1007/s11097-007-9076-9;
- De Jaegher, H, Pieper, B., Clénin, D., & Fuchs, T. (2017) Grasping intersubjectivity: an invitation to embody social interaction research, *Phenomenology and the Cognitive Sciences*, 16(3), 491-523. doi: 10.1007/s11097-016-9469-8;
- Eilan, N., Hoerl C & Roessler, J. (2005) (eds) *Joint Attention: Communication and Other Minds*. Oxford: Oxford University Press;
- Eilan, N. (ed.) (2014) ‘The Second Person’ Special Issue, *Philosophical Explorations*, 17(3), 265-387;
- Ferencz-Flatz, Christian (2019) Ten theses on the reality of video-chat: A phenomenological account, *Communications*, 44(2), 204-224. doi: 0.1515/commun-2018-2012;
- Fivaz-Depeursinge, E. & Corboz-Warnery, A. (1999) *The Primary Triangle*. New York: Basic Books;
- Fuchs, T. (2014) The Virtual Other: Empathy in the Age of Virtuality, *Journal of Consciousness Studies*, 21(5-6), 152-173
- Fuchs, T. (2017) Intercorporeality and Interaffectivity. In Meyer, C; Streeck, J & Jordan, J. Scott, *Intercorporeality: Emerging Socialities in Interaction* (pp. 3-24). New York: Oxford University Press;
- Gallagher, S. (2014) ‘Phenomenology and Embodied Cognition’ in L. Shapiro (ed.) *The Routledge Handbook of Embodied Cognition* (pp. 9-18). London; New York: Routledge;
- Gallagher, S. (2017) *Enactivist Interventions: Rethinking the Mind*. London: Oxford University Press;
- Gibson, J.J. (1986/2015) *The Ecological Approach to Visual Perception*. New York: Routledge;
- Gilbert, M. (2014) *Joint Commitment: How We Make the Social World*. New York: Oxford University Press;
- Hobson, P. (1989) On sharing experiences, *Development and Psychopathology*, 1, 97-203. doi: 10.1017/S095457940000390
- Hobson, P. (2008) Interpersonally Situated Cognition, *International Journal of Philosophical Studies* 16(3), 377-397. doi: 10.1080/09672550802113300
- Krueger, J. & Osler, L. (2019) Engineering Affect: Emotion Regulation, the Internet, and the Techno-Social Niche, *Philosophical Topics*, 47(2), 205-231, doi:10.2307/26948114;
- Merleau-Ponty (1945/2012) *Phenomenology of Perception*, trans. D.A. Landes. London; New York: Routledge;
- Morganti, F., Carassa, A. & Riva, G. (2008) *Enacting Intersubjectivity*. Amsterdam: IOS Press;

- Morganti, F. (2008) What intersubjectivity affords: paving the way for a dialogue between cognitive science, social cognition and neuroscience (pp3-16). In Morganti *et al* (2008);
- Newen, A., De Bruin, L. & Gallagher, S. (eds.) (2018) *The Oxford Handbook of 4E Cognition*. Oxford: Oxford University Press;
- Palmer, J. (2019, August 30) Spatial Interfaces. [Web log post]. Retrieved from: <https://darkblueheaven.com/spatialinterfaces/>
- Palmer, J. (2020, April 9) Spatial Software. [Web log post]. Retrieved from: <https://darkblueheaven.com/spatialsoftware/>
- Pawlett Jackson, S. (2018) Gestalt structures in multi-person intersubjectivity, *Synthese*. doi: 10.1007/s11229-018-02001-y
- Pawlett Jackson, S. (2019) Exploring different intersubjective structures in relation to dialogue. *Arts and Humanities in Higher Education*. 18(1) 22-33;
- Osler, L. (Forthcoming) 'Taking Empathy Online', *Inquiry*.
- Reddy, V. (2005) Before the Third Element: Understanding Attention to Self (pp. 85-109). In Eilan (2005);;
- Reddy, V. (2008) *How Infants Know Minds*. London: Harvard University Press;
- Reddy, V. (2015) Joining intentions in infancy, *Journal of Consciousness Studies*, 22(1-2), 24-44.
- Seemann, A. (2012) *Joint Attention: New Developments in Psychology, Philosophy of Mind, and Social Neuroscience*. Cambridge, Mass: The MIT Press;
- Stawarska, B. (2009) *Between You and I: Dialogical Phenomenology*. Athens, Ohio: Ohio University Press;
- Varela, F., Thompson, E. & Rosch, E. (1991) *The Embodied Mind: Cognitive Science and Human Experience*. USA: The MIT Press;
- Zahavi, D. (2014) *Self and Other*. New York; Oxford: Oxford University Press;
- Zahavi, D. (2015) You, Me, and We: The Sharing of Emotional Experiences. *Journal of Consciousness Studies*, 22(1-2), 84-101.