A Multiplicity View for Social Cognition: Defending a Coherent Framework

A Reply to Lisa Quadt

Albert Newen

Lisa Quadt's commentary focuses on my theory about the multiple epistemic strategies humans use to receive information about one other's mental phenomena. She develops a principle worry about the theory's underlying metaphysical foundations, arguing that I am committed to an incoherent metaphysical framework. In this reply, I show that I am not committed to the position she attributes to me and I outline an alternative framework that is my actual background view. I illustrate this framework by discussing emotions and argue that emotions are individuated as integrated patterns of characteristic features. This enables me to combine a representational account of emotions with a theory of direct perception of basic emotions as well as with an understanding of some emotions relying on theory-based inferences. Thus, I have a coherent metaphysics. Finally, I show that the alternative suggested by Quadt has its own problems.

Keywords
Direct perception | Metaphysical foundation | Person model theory | Social cognition | Transparency

Author

Albert Newen
albert.newen@rub.de
Ruhr-Universität Bochum
Bochum, Germany

Commentator

Lisa Quadt
lisquadt@students.uni-mainz.de
Johannes Gutenberg-Universität
Mainz, Germany

1 Introduction

With my PMT (person model theory), I aim to answer two questions. While the first question asks which epistemic strategy humans use to access the mental states of others and to gather information about them, the second question asks how the information we obtain to understand others is stored and organized. The answer to the second question is the core of the PMT. It states that information about other individuals or types of persons is stored and organized in person models and that these are realized on two levels, i.e. the implicit level of person schemata and the explicit level of person images. It further argues that philosophical theories so far have predominately ignored the fact that we usually understand others relying on rich background information concerning them and their situation.

Lisa Quadt's commentary focuses on my theory concerning the epistemic strategies humans use to receive information about others' mental phenomena, and she develops a principle worry about the underlying metaphysical foundations. I am grateful for this challenge, which gives me the opportunity to clarify my background view. The MV (multiplicity view) outlined in the target paper claims that we do not rely on one epistemic strategy alone, as is suggested by most proposals in the literature, but that we rely on a multiplicity of strategies which, for the most part, are implicitly activated on the basis of contextual conditions. These strategies include simulation strategies, theory-based inferences, and direct perception as well as understanding by social interaction and by relying on narratives. Quadt's main worry is that MV may be based on an incoherent metaphysics and is thus inacceptable as it stands. In the first part of her reply she aims to defend the incoherence claim, while in the second part she offers an alternative metaphysical framework. My reply is structured as follows: In the next paragraph I briefly describe how Quadt defends her claim about the supposed incoherence of my metaphysical background and show that I am not committed to the incoherent framework she attributes to me. In the second section, I make explicit my actual background metaphysics (which was not the focus of my article) and argue that it is coherent, reinforcing that I am not committed to the metaphysics that Quadt attributes to my position. Finally, I argue that the alternative metaphysics suggested by Quadt relies on a distinction between transparency and opacity that cannot carry the weight it is supposed to carry.

2 Am I committed to an incoherent metaphysics?

Quadt describes correctly that the MV I advocate combines epistemic strategies that are described in several different positions, including ST (Simulation Theory) (Goldman 2006), TT (Theory-Theory) (e.g., Gopnik & Meltzoff 1997), and IT (Interaction Theory) (Gallagher 2001), as well as theory of direct perception (Gallagher 2008). As a consequence, she presupposes that I am committed to the metaphysical foundations of each of these positions, while each position argues for a distinct epistemic strategy. If I were committed to accepting such metaphysical foundations, I would thereby offer an incoherent metaphysics. Quadt shows this by arguing that Simulation Theory and Theory-Theory, on the one hand, presuppose metaphysical claims that are not consistent with the presuppositions from Theories of DP (direct perception) and ITs, on the other hand (3). Quadt claims that ST and TT are *cognitivist theories* that presuppose internalism, mental representations, and the idea that mental phenomena are private hidden entities to which we have no direct access. To register mental phenomena we have to rely on perceiving the behaviour and expressions of other people and have to *infer* the existence of mental phenomena. Quite the opposite view is taken by the *non-cognitivist theories* of DP and IT. They allow for externalism of mental phenomena (as being realized by two people and their interaction), they deny the existence of mental representations, and they presuppose that mental phenomena are not hidden but directly perceivable. Thus they rely on non-inferential access to mental phenomena by direct perception. The following quote illustrates the main features of the contrast Quadt develops:

> The difference between cognitivist and non-cognitivist pictures of social cognition, in the cases that I just described, seems to boil down to the metaphysical assumption of whether or not there are hidden cause in the outside world that require an inference or representational mechanism in order to access and process them. While ST and TT clearly assume such a view, DP denies it. Therefore, I claim that MV cannot simply combine theoretical elements that draw on such considerable metaphysical differences. (Quadt 2015, p. 5)

My first general reply to this worry is that I only take on the description of an *epistemic strategy* of acquiring and using information

about other people in order to understand them. An epistemic strategy like a simulation (to put oneself in the other person's shoes) or a theory-based inference is not automatically connected to a metaphysical commitment. De facto, the philosophers who are famous for holding ST or TT combine their view with a metaphysical background, but it does not follow that the epistemic strategy they describe *must be combined* with the metaphysical background they offer. We can easily see this for example in the case of two epistemic strategies like theory-based inferences and direct perception of mental phenomena. These can be easily combined in a way that allows that some mental phenomena with intense expressive components like basic emotions (Ekman et al. 1972) can be directly perceived (see below), while complex mental phenomena like propositional attitudes may be at least often inferred if the social understanding cannot rely on honest utterances but only on some ambiguous behavioural cues. Thus, the de facto incompatibility of the metaphysical presuppositions of the two main lines of theories of social understanding does not imply that I am committed to inheriting both presuppositions and that I thus run into an incoherent metaphysics. In fact, I do not presuppose two metaphysical principles for the same mental phenomenon; instead I only need to allow for the application of two epistemic strategies of understanding mental phenomena, which may be applied to different mental phenomena (or to the same type of mental phenomenon in different situations). In the next section I outline my alternative metaphysics and illustrate both that it is coherent and that it can allow for direct perception as one epistemic strategy for registering some mental phenomena.

3 Defending direct perception in an alternative metaphysical framework

In general, I prefer to think of mental phenomena as representational, but I do not see that this prevents me from integrating the epistemic strategy of direct perception. Furthermore, I characterize basic emotions as realized in one individual (individualism but not internalism). At the same time, I remain neutral as to whether joint emotions (e.g. joint enthusiasm about a goal achieved by one's team) have to be analysed as extended emotions. Furthermore, I think that basic emotions are not hidden mental phenomena but can be directly perceived e.g. on the basis of face-based recognition of emotions. Thus, I think that some mental phenomena can be registered non-inferentially. But of course, direct perception of some mental phenomena is *only one* of at least four epistemic strategies that we can use, depending on the context.

To sketch my theory of direct perception I will focus on basic emotions like anger, fear, happiness, sadness, etc. (for a classification of emotions see Zinck & Newen 2008). My metaphysical view of emotional episodes is that they are integrated patterns of characteristic features (Welpinghus & Newen 2012; Newen et al. 2015). Let me use the example of fear as illustrated in Newen et al. (2015): an emotional episode of fear towards an aggressive dog is constituted by the integration of the following characteristic features: (1) a typical physiological arousal that is a consequence of bodily changes due to changes in the autonomic nervous system, including increased heart rate and flat breathing; (2) a typical behavior or behavioral disposition, including flight or freezing behavior; (3) a typical facial expression, gesture, or body posture, etc.; (4) a typical phenomenal experience of fear; (5) a typical (explicit) cognitive evaluation of the dog in front of me (e.g., "This is an aggressive pit bull"). Furthermore, every emotional episode has (6) an intentional object, i.e. the dog in front of me. Features 1–5 are integrated into an (often implicit) appraisal of the intentional object as dangerous. The emotional episode is constituted by the integration of all the characteristic features mentioned so far, including the appraisal. This view allows that in another implementation *some* features would be missing. For example, the explicit cognitive evaluation of the dog as an aggressive pit bull is not necessary to be in fear towards the dog in front of me. Or the facial expression may be inhibited, due to intense training to attain a poker face, yet I may still be in fear. As long as a minimum of features is realized, we still have

an episode of fear. The two main features that are necessary in all emotional episodes are a registration of minimal physiological arousal and an intentional object. The integration of both is needed to have an emotional episode (Barlassina & Newen 2013). But other features may be lacking while still remaining characteristic of most episodes of the relevant type of emotion. One might wonder why I do not include neural correlates. Since I argue from a position of antecedent naturalism, neural correlates are not an extra component in addition to the characteristic features already mentioned above. We might mention neural correlates as an informative aspect for the individuation of certain features of emotion, but we do not have to, since they concern the same features that have already been mentioned, with information accessed in a different manner.

If one accepts the ontology of emotions as individuated by an integrated pattern of characteristic features, it follows that the expression of an emotion by face, body posture, and gestures is a *constitutive* part of the emotional episode (and not a causal consequence). Thus, I do not hold internalism about mental phenomena. Given this theory of the individuation of emotions, I also argue for the thesis that one way of recognizing emotions is by perceiving the relevant pattern (Newen et al. 2015). A recognition of the other person's fear can be attained by directly perceiving the pattern of fear. How can we account for this, while at the same time accepting that the feeling of fear is a private subjective experience in so far as a person still may have the feeling even if she is able to keep a poker face? Perceiving fear is comparable to perceiving a house. Both are processes of pattern recognition on the basis of a minimal package of characteristic features: I can recognize a drawing as one of a house, even if one or two of the characteristic features of a house are missing. How is this possible? Perceiving an object is not a purely passive process, like taking a photograph; it is a constructive process.[1] One important aspect of the constructive process is the enrichment of selected core sensory information. And one way of realizing this enrichment is by the activation of a rich memorized mental image that best suits the core sensory information. If we have learned the relevant pattern of what a house looks like from the outside, and memorized a respective mental image, then seeing a child's drawing initiates an interaction of bottom-up and top-down processes. These include the activation of this stored mental image, such that it enriches the core sensory information to form a perceptual experience of seeing a drawing of a house even if the front door is missing in the drawing.

The same process of pattern recognition takes place in the case of recognizing an emotion like fear. The relevant pattern of fear is formed either on the basis of having personally experienced a situation of fear or on the basis of having observed others in such situations. One thereby acquires a memorized pattern of fear with typical features. If one now observes a person with a typical facial expression in a situation where she is being attacked by a dog, one can see the fear of the person. The perception of fear is realized by seeing the freezing behaviour, the facial expression, and the intentional object (i.e. the aggressive dog), because these features activate as part of the process of perceptual processing the whole pattern of fear. Thus, I can *perceive* fear in the face of the person being attacked by the dog. The theory of perception is one according to which perceptual processing allows for a systematic enrichment of information and for influencing of perceptual processes by memorized images or background knowledge. These top-down influences are discussed under the label cognitive penetration. So I am committed to the view of perception as cognitively penetrated as it is defended in detail in Vetter & Newen (2014). But this does not involve any claims concerning the metaphysical commitments ascribed to me by Quadt in her commentary. Recognition of emotions is analysed in a framework that explicitly allows for mental representations but specifies them in a way that nevertheless allows for direct perception as one form of access to the recognition of

1 All modern theories of perception account for this constructive component, e.g. O'Regan's and Noë's theories of enacted perception (O'Regan & Noë 2001; Noë 2005), as do theories of cognitive penetration (Macpherson 2012; Siegel 2012) and theories of predictive coding (Hohwy 2013; see alsoHohwy this collection; Clark this collection).

emotions. As has been spelled out in detail elsewhere (see Newen et al. 2015), in principle I allow for three types of recognizing of emotions: two types of direct perception are distinguished in terms of top-down processes of shaping perception involving background images or beliefs; and one is characterized by theory-based inferences. Thus, I distinguish "(1) (a basic form of) perceiving an emotion in the (near) absence of any top-down processes, and (2) perceiving an emotion in a way that significantly involves some top-down processes (a strongly concept-modified form of perception). Both types of perceiving emotions can be distinguished from (3) inference-based evaluation of an emotion pattern. The latter presupposes a stable evaluation of an emotion as being F, which then may be modified or reevaluated by reflecting on the information" (Newen et al. 2015, p. 197). To sum up: Direct perception can be based on a metaphysical framework that regards emotions as integrated patterns of characteristic features and this allows me to combine it with presupposing mental representations of emotions (as memorized rich patterns), on the one hand, as well as with a non-inferential recognition of some emotional episodes on the other. The pattern theory of emotion is furthermore able to account for internalistic features of emotions like the feeling of fear, but also for individualistic yet expressive features like behavior and expression in face, gesture, and body posture. This metaphysics of emotions is coherent and is compatible with several epistemic strategies for recognizing them, e.g. direct perception as well as theory-based inferential understanding.

Let me make a further clarificatory remark about my reply to the coherence worry: I illustrated my metaphysics taking emotional episodes as a core example. This does not imply that I analyze *all* mental phenomena in this way. Although I think that some mental phenomena can also be individuated as integrated patterns of characteristic features like self-awareness/self-consciousness (see Gallagher 2013) or object perception, I remain neutral on the question of how far this analysis can be generalized and about the possibility that some mental phenomena need a different metaphysics as basis. For this reply it is sufficient to have shown what a concrete paradigmatic example of a coherent metaphysics for emotional episodes looks like, in order to prevent the danger of running into an incoherent metaphysics as a unavoidable consequence of the multiplicity view concerning epistemic strategies of understanding others.[2]

4 Quadt's proposal FOR an alternative metaphysical framework

Although I think I do not need an alternative metaphysics, since I have a coherent one already, I would like to briefly comment on Quadt's account. She starts with a remark on embodiment. I do not really see any serious disagreement with my views here. For it is fine by me that phenomenal properties and mental representations in general are realized within the body —and sometimes not only in the brain but within our whole body (see the discussion of emotions). Furthermore, I said that in this reply I leave open whether we need an extended realization basis for some mental representations. Quadt's alternative proposal, with which she aims to deliver a new framework for a multiplicity view, introduces different levels of embodiment. One way to read her distinction is that it offers a characterization of different types of representation that unfold during ontogeny. This basic idea is entirely consistent with my work. In other papers I discuss in detail the development of different types of representation in ontogeny (Newen & Vogeley 2003; Newen & Fiebich 2009; de Bruin & Newen 2012). There are of course differences in how one might form types of representation but discussion of these goes beyond the scope of this reply.

Let me now elaborate on an important point of disagreement. Quadt's proposal is based, among other things, on the distinction between transparent and opaque ways of being involved in a mental state. She takes this distinction from Metzinger (2003, 2004). We can illustrate this dis-

2 Let me highlight that the multiplicity view of understanding others is only one part of my person model theory and this epistemic aspect is in addition defended and further developed by my former PhD-student Anika Fiebich in the following paper which just appeared: Fiebich & Coltheart 2015.

tinction using the example of the mental event of perceiving an apple. This event is transparent if I am only consciously aware of the apple, while it is opaque if I am (also) aware of my mental state of seeing the apple: "[w]hat distinguishes transparent from opaque states is the degree to which one's own social cognitive processing, which is directed at the other person, is explicitly represented as a process" (Quadt 2015, p. 12). The relevant move is Quadt's claim that the epistemic access of direct perception in social cognition can be explained by transparency, while the epistemic access of simulation and theory-based inference can be explained by opacity.

Here I think she is on the wrong track. This distinction between transparency and opacity in the case of a mental state of attributing a belief leads to the idea that I am not only aware of the other person having a belief with content p but that I am also focussing on being consciously aware of the process of my attributing a belief to the other. The latter can of course happen in case of reflective processes of attributing beliefs; but normally we are in a mode of just using our ability to attribute beliefs automatically, focusing on the other's belief and its content (not on our own process of attributing it). We normally deal with our mental state of attributing beliefs in a transparent way, contrary to the analysis offered by Quadt. Furthermore, direct perception can also be used opaquely in rare cases of being reflectively aware of guiding images: if I am an experienced chess player, I can perceive the chess board in a way that is best described by cognitive penetration, and in some cases I may be aware of the mental image which guides my perception, i.e. I see a position and know how to act because I consciously memorize the fact that I see exactly the same position I saw in a previously played game. Thus, the distinction between transparency and opacity is not helpful for characterizing the different strategies of epistemic access to another's mental states.

5 Self-models and person models: how are they related?

Finally let me point out an important question raised by Quadt, namely how are person models and self-models related to each other? A self-model is a special type of person model, the person model that someone develops of herself. This is also done at the two levels of an implicit self-schema and an explicit self-image. I intend to elaborate on the interaction between self-models and person model of others in future articles, but I completely agree with Quadt when she says that there is bi-directional informational exchange regarding both types of models in humans (which is also indicated in my paper in figure 2, p. 21): "I thus conclude that it should not only be considered how the development of a self-model influences social cognition, but also which role social processes play in forming such a self-model" (Quadt 2015, p. 10). The PMT has potential as a framework for a theory of human self-consciousness.

References

Barlassina, L. & Newen, A. (2013). The role of bodily perception in emotion: In defense of an impure somatic theory. *Philosophy and Phenomenological Research*, 1-42.

Clark, A. (2015). Embodied prediction. In T. Metzinger & J. M. Windt (Eds.) *Open MIND*. Frankfurt a. M., GER: MIND Group.

de Bruin, L. & Newen, A. (2012). An association account of false belief understanding. *Cognition*, *123* (2), 240-259. 10.1016/j.cognition.2011.12.016

Ekman, P., Friesen, W. V. & Ellsworth, P. (1972). *Emotion in the Human Face.* Oxford: Pergamon Press.

Fiebich, A. & Coltheart, M. (2015). Various ways to understand other minds. *Mind and Language.*

Gallagher, S. (2001). The practice of mind: Theory, simulation, or interaction? *Journal of Consciousness Studies*, *8* (5-7), 83-107.

——— (2008). Direct perception in the intersubjective context. *Consciousness and Cognition*, *17* (2), 535-543. 10.1016/j.concog.2008.03.003

——— (2013). A pattern theory of self. *Frontiers in Human Neuroscience*, *7* (443), 1-7.

Goldman, A. I. (2006). *Simulating minds: The philosophy, psychology, and neuroscience of mindreading.* New York, NY: Oxford University Press.

Gopnik, A. & Meltzoff, A. N. (1997). *Words, Thoughts, and Theories.* Cambridge, MA: MIT Press.

Hohwy, J. (2013). *The Predictive Mind.* Oxford: Oxford UP.

——— (2015). The neural organ explains the mind. In T. Metzinger & J. M. Windt (Eds.) *Open MIND*. Frankfurt a. M., GER: MIND Group.

Macpherson, F. (2012). Cognitive penetration of colour experience. Rethinking the issue in light of an indirect mechanism. *Philosophy and Phenomenological Research*, *84* (1), 24-62. 10.1111/j.1933-1592.2010.00481.x

Metzinger, T. (2003). Phenomenal transparency and cognitive self-reference. *Phenomenology and the Cognitive Sciences*, *2*, 353-393.

——— (2004). *Being no one: The self-model theory of subjectivity.* Cambridge, MA: MIT Press.

Newen, A., Welpinghus, A. & Juckel, G. (2015). Emotion recognition as pattern recognition: the relevance of perception. *Mind & Language*, *30* (2), 187-208.

Newen, A. & Fiebich, A. (2009). A developmental theory of self-models. In W. Mack & G. Reuter (Eds.) *Social Roots of Self-Consciousness. Psychological and Philosophical Contributions* (pp. 161-186). Berlin, GER: Akademie 2009.

Newen, A. & Vogeley, K. (2003). Self-representation: searching for a neural signature of self-consciousness. *Consciousness & Cognition*, *12* (4), 529-543. 10.1016/S1053-8100(03)00080-1

Noë, A. (2005). *Action in Perception.* Cambridge, MA: MIT Press.

O'Regan, J. K. & Noë, A. (2001). A sensorimotor account of vision and visual consciousness. *Behavioral and Brain Sciences*, *24* (5), 939-1031.

Quadt, L. (2015). Multiplicity Needs Coherence – Towards a Unifying Framework for Social Understanding. In T. Metzinger & J. M. Windt (Eds.) *Open MIND*. Frankfurt a. M., GER: MIND Group.

Siegel, S. (2012). Cognitive Penetrability and Perceptual Justification. *Nous*, *46* (2), 201-222.

Vetter, P. & Newen, A. (2014). Varities of cognitive penetration in visual perception. *Consciousness and Cognition*, *27*, 62-75.

Welpinghus, A. & Newen, A. (2012). Emotion und Kultur. Wie individuieren wir Emotionen und welche Rolle spielen kulturelle Faktoren dabei? *Zeitschrift für philosophische Forschung*, *66* (3), 367-392.

Zinck, A. & Newen, A. (2008). Classifying emotion: a developmental account. *Synthese*, *161*, 1-25.