e-ISSN 2723-9640

JoLMA
Vol. 6 — Num. 1 — June 2025

Why Direct Social Perception Theory Needs To Be More Gibsonian

Martyna Meyer
University of Southern Denmark, Denmark

Edward Baggs
University of Southern Denmark, Denmark

Abstract The current debate on social perception is dominated by the invisibility principle: the assumption that mental states are inherently private and unobservable. Direct social perception theorists challenge this view, arguing that the mental states of others can be perceived directly. At present, however, there is no widely agreed programmatic approach for studying social perception as a direct phenomenon. In this paper, we identify three strategies that have been proposed by researchers to advance the direct social perception account. The empirical strategy emphasizes the need for experimental evidence; the Anscombean strategy highlights the importance of separating the analyst's description of an action from the actor's own perspective; finally, the Gestalt strategy proposes viewing mental states, such as emotions, as patterns of structures that occur across space and time. We argue that all three strategies can be pursued within the Gibsonian framework. The Gibsonian approach stresses the importance of ecological information, understood as structures in the environment discriminable in perception. We suggest that the Gibsonian perspective offers a coherent foundation for understanding social perception without relying on mentalistic inference.

Keywords Direct Social Perception. Ecological Information. Gestalt. Anscombe. Ecological Psychology.

Summary 1 Introduction. – 2 Direct Social Perception and Embodied Cognitive Science. – 3 The Empirical Strategy. – 4 The Anscombean Strategy. – 5 The Gestalt Strategy. – 6 The Gibsonian Framework: Information for Direct Social Perception.

Edizioni Ca'Foscari

Peer review

Submitted 2025-01-08
Accepted 2025-03-07
Published 2025-06-13

Open access

Citation Meyer, Martyna; Baggs, Edward (2025). "Why Direct Social Perception Theory Needs To Be More Gibsonian". *JoLMA*, 6(1), 31-50.

DOI 10.30687/Jolma/2723-9640/2025/01/002

1 Introduction

Do we ever perceive another person's anger? Certainly, we can perceive the other person's facial expressions and their behavior. According to a common intuition, however, this is not the same thing as saying that we perceive their anger: what we perceive is simply the outward manifestation of the emotion; there is something inherently subjective about the state of being angry. We can push this intuition a little further. Varga (2020) proposes the following thought experiment. Imagine we had access to a machine that could show us in real time the neural makeup of the person undergoing an episode of anger. Would we then be inclined to say that we see the person's anger? Varga (2020) suggests that we would be equally inclined to say 'no'. We would say that we can see the brain scan of an angry person, but still not the anger itself.

This thought experiment brings out an important assumption about what mental states are: we tend to see them as inherently undiscoverable and unobservable. Even advanced screening technology would not be able to make us see emotions, because emotions are private, inaccessible to others: they exist within subjective experience. Fernández Castro and Heras-Escribano (2020) call this view 'the assumption of invisibility'. It is also known as 'the unobservability principle' (e.g., Krueger 2012; Becchio et al. 2018). According to this assumption, certain parts of the world are not available to be seen, but can only be inferred. The assumption of invisibility belies a deeply dualist understanding of the mental.

A number of theorists have sought to deny the assumption of invisibility, and have proposed that perception of mental states can be direct. To date, however, no coherent positive programmatic approach to direct social perception has been developed. The debate remains overly conceptual and somewhat confused. The paper is structured as follows. We provide a brief overview of the state of the debate. We then individuate three methodological strategies that have recently been proposed for addressing these issues: the empirical strategy, the Anscombean strategy, and the Gestalt strategy. We will conclude by suggesting that all three of these strategies can best be pursued within the framework of Gibsonian perceptual information.

2 Direct Social Perception and Embodied Cognitive Science

The notion of 'direct social perception' has been the subject of some debate among philosophers since the late 2000s (Gallagher 2008; Krueger 2018; Spaulding 2015; Zahavi 2011). The driving force of this debate was, initially, dissatisfaction with the theory of mind paradigm. The theory of mind paradigm itself arose out of Premack and Woodruff's (1978) classic study asking, "Does the chimpanzee have a theory of mind?" Daniel Dennett's commentary on the paper, which offered a critique of the methodology, led to subsequent studies of human children with a focus on whether children are sensitive to the false beliefs of others (Dennett 1978; Wimmer, Perner 1983). Much of the research on social cognition in the 1980s and 1990s centered around the question of how we come to know others' mental states (including their false beliefs). For many years, the two main hypotheses were the theory-theory which claimed that we learn the landscape of others' mental states by hypothesis-building and -testing, much like scientists discovering the nature of the world (Gopnik et al. 1999), and the simulation theory, which claimed that we grasp others' mental states by implicitly or explicitly simulating them on our own neural apparatus (Gallese, Goldman 1998; Gordon 1986).

Neither of these hypotheses is particularly attractive from the viewpoint of embodied cognitive science. Both posit social cognition as a private achievement of the observer. The most prominent alternative hypothesis to the theory of mind paradigm is Shaun Gallagher's interaction theory (Gallagher 2004; 2008; 2020). In the most comprehensive account of this theory to date, Gallagher (2020) builds a negative argument against the theory of mind paradigm. He usefully identifies not one but three problematic assumptions associated with theory of mind (71). These are:

1. the unobservability principle, i.e. the assumption that we don't have direct perceptual access to other minds
2. the observational stance, i.e. the assumption that we know about others by observing them
3. the supposition of universality, i.e. the assumption that a theory of mind mechanism must explain all instances in which we understand others

The first assumption we have already discussed above. Gallagher proposes to reject all three assumptions. He asserts instead (Gallagher 2020, 100) that (1) other minds are not hidden away and inaccessible, (2) our normal everyday stance towards others is second-person, not third-person, and (3) our primary way of understanding others is not through mindreading but through embodied engagement.

Similarly, Fernández Castro and Heras-Escribano (2020) identify two problematic assumptions within existing work on social cognition. First is the aforementioned assumption of invisibility, which holds that mental states are inherently private to the individual who undergoes them. Second is the assumption of methodological individualism, according to which the primary unit of analysis should be the cognitive states of an individual cognizer, as opposed to, for example, the interacting dyad or the entire ecological, cultural and social environment of the agent.

This second assumption has been the focus of much enactivist theorizing about social cognition, notably the theory of participatory sense-making (De Jaegher, Di Paolo 2007; Di Paolo et al. 2018). This refers to collaborative activity within which the agents actively regulate their coupling with each other and with the environment, giving rise to an emergent whole that no individual is in control of (think of the movement of players on a soccer team as they create and exploit space around their opponents). The agent and the environment are dependent on and shaped by one another. Arguably, the entire agent-social environment pair is the unit on which level cognition takes place (De Jaegher et al. 2010).

But let us return specifically to the notion of direct social perception. How should we understand it? How does it differ from indirect social perception? Gallagher (2008) explains the difference in a useful way. Direct perception, he suggests, is smart; indirect perception is dumb. The example Gallagher gives is that when looking at a car, the dumb perception account posits that we perceive components (redness, shapeness, etc.) which we then mentally combine 'to make it add up to my car' (Gallagher 2008, 536). The smart perception account, meanwhile, posits that we perceive the car as a whole. The entire configuration is available for perceptual detection.

> The question about direct perception, whether of objects or of others is not simply about how direct it is, or what directness means, but how smart, how richly informed, it is. The smarter the perception is, the more work it does; the dumber it is, the more it requires extra cognitive processes (theory, simulation) to get the job done. The direct perception theorist is claiming that social perception is very smart and that in the usual circumstances of social interaction it does most of the work without the need of extra cognitive (theoretical or simulationist) processes. (Gallagher 2008, 538)

In our opinion, the distinction Gallagher draws in the above quote is key to any coherent understanding of the phrase 'direct social perception.' At the same time, the approach that Gallagher is advocating here appears to be an instance of adopting 'the observational stance,' which, in the later text, he claims is problematic (Gallagher

2020). We know about the car, in Gallagher's example, by observing it. That is, we know about it by interacting with the information that it projects into light, not by interacting physically with the car. Gallagher's (2008) notion of direct social perception as smart perception seems to be separable from the notion that social cognition is inherently interactive (in the social sense), and yet, as can be seen from the above, these two issues have been conflated in the literature on embodied social cognition (De Jaegher 2009).[1]

In summary, the habitual framing of embodied accounts of social cognition as being a response to, or an alternative to, the theory of mind paradigm appears to have led to a couple of undesired consequences. First, embodied accounts of social cognition have tended to become overly negative, focusing on what theory of mind gets wrong instead of how to build a coherent programmatic alternative. What is called for is a positive theory of social cognition within the embodied tradition. Second, embodied accounts have tended to conflate two issues: on the one hand, the interactive, or second-person, nature of social life, and, on the other, the availability of information about other actors in the environment.

In the rest of this paper, we will focus on some desiderata for a positive programmatic approach to direct social perception.

3 The Empirical Strategy

If our aim is to develop a programmatic approach to the perception of social properties, a reasonable place to start is to ask what kind of empirical evidence we should be seeking. We take our cue here from a recent discussion by Mason Westfall (2023) on the perception of agency. Westfall argues that the perception of agency is fundamentally the same thing as the perception of animacy. To empirically investigate agency, therefore, a good place to start would be to consult the literature on the perception of animacy. As Westfall points out, however, existing research from within the theory of mind paradigm is not necessarily helpful here. 'Most discussions of animacy center on the role animacy plays in mental state attribution–in particular, the attribution of intentions. Often, though, we perceive animacy *without* attributing mental states' (Westfall 2023, 848; italics in original).

1 In the dynamical systems tradition, some theorists attempt to understand social interaction as arising from spontaneous interpersonal coordination (e.g. Richardson et al 2007). Again, we understand this work as being about processes of social interaction, and somewhat separate from the direct social perception discussion, which, we are suggesting, should be concerned with identifying information or perceptual invariants. Thanks to an anonymous reviewer for encouraging us to address this line of work explicitly.

Westfall (2023) argues that perceiving animacy is not the same as perceiving motion, because it is possible to perceive something as animate without perceiving it moving, e.g. a motionless deer encountered on a hiking trail. It is also not the same as perceiving unpredictable motion, because even unpredictably moving entities such as a plastic bag in the wind do not appear animate. Finally, perceiving animacy is not the same as perceiving life, because we can distinguish life from animacy, and animacy detection is developmentally prior to life detection.

The most famous empirical study of the perception of animacy is Fritz Heider and Marianne Simmel's classic study of 'apparent Behavior' (Heider, Simmel 1944). In the study, the subjects are presented an approximately 2.5 minute long film in which two triangles and a circle move around a rectangular, empty structure. After seeing the film, the subjects are asked to interpret what happened. Subjects tend to spontaneously describe the figures as intentional entities, having personalities and complex goals. The Heider and Simmel study is an excellent example of the human tendency to see intentionality in seemingly the most simple structures.

While the Heider and Simmel study is compelling as a demonstration that we can perceive both animacy and agency in simple motion, the study does not establish how we achieve this perceptual outcome. What are the underlying mechanisms that lead participants to see animacy when they watch the film? The theory of mind paradigm reaches for an explanation in terms of ascription: participants perceive animacy and intentions because they ascribe personalities to the figures. What we are interested in, however, is whether there is a perceptual basis for the appearance of animacy. Scholl and Gao (2013) provide a useful set of clarifications here:

> "[P]erception" refers in this context to a family of processes that is relatively *automatic* and *irresistible*, and that operates *without the ability to consciously introspect its nature* [...] for a process to count as perceptual, it needs to be *strongly and directly controlled by specific and subtle features of the visual input itself*. (Scholl, Gao 2013, 202; italics added)

Note the compatibility with Gallagher's (2008) definition of direct perception as smart perception, discussed above. A perceptual account of animacy-perception, on this reading, would be one that does not require the observer to add something mental on top of the incoming stimulus information.

An example of a class of processes that can be understood as perceptual, in the sense described here, is the perception of causation. Causal perception is retinotopically specific and cannot be explained by appeal to adaptation of lower-level visual features (Rolfs et al.

2013). It would be very difficult to explain the perception of causation without an appeal to visual processing because we do not know of any higher-level processes (like cognition or judgments) that evoke retinotopically specific processing. This is the standard that we would like animacy to meet. Scholl and Gao (2013) suggest, pessimistically, that this particular retinotopically specific processing goal cannot be achieved by animacy because the perception of animacy is dynamic in both spatial and temporal dimensions.

The perception of animacy has been operationalized in a number of different ways, for example: as self- propulsion, as abrupt changes in speed or direction, as certain patterns of approach and avoidance, as coordinated orientation cues. Generally, animacy is thought to be perceived when agents violate Newtonian mechanics. Animate objects move in ways that cannot be explained by the forces their environment subjects them to (Scholl, Gao 2013). To argue that animacy is directly perceived we would have to gather evidence that self-propelled motion is distinguished from other stimuli by a specific bottom-up visual processing mechanism.

Westfall (2023) discusses multiple studies that he argues provide convergent evidence for a perceptual understanding of animacy-detection. We focus on one of these. A study by Pratt et al. (2010) uses a multiple object tracking paradigm to investigate whether attention is preferentially, and automatically, allocated to animate motion. The authors conducted a series of experiments that suggest that subjects perform better at detecting changes in geometrical objects if these objects have previously exhibited animate movement. In the study, animacy is defined as self-produced and unpredictable motion.

The subjects in Pratt et al.'s (2010) study see a number of geometrical figures on a screen. The figures that exhibit unpredictable motion, that is, those which show changes in motion that are not a result of collisions, are considered to be the 'animate' objects. In the first two experiments, four objects move pseudo-randomly on the screen, sometimes colliding with each other and with the screen frame. At one point, one of the objects changes direction without a preceding collision. Subsequently, one of the four figures changes its state and the participants need to report this change. The hypothesis is that if the changed object was the one that exhibited animate motion, the reaction of the study participants will be faster, since their attention is already allocated to this object. In both experiments, the subjects were indeed faster to respond to the change if it involved the object that exhibited animate motion.[2]

2 These two experiments are followed-up by four additional experiments, excluding that the faster reaction time was due to the animate movement being a singleton (an odd one out) or that the subjects reacted faster because they learned to strategically

We will not here attempt to resolve the issue of whether there can be a purely perceptual explanation of the detection of animacy. Interested readers are directed to the discussion in Westfall (2023). The moral that we want to draw is that whether or not something is perceptual is an empirical matter. The problem with the theory of mind paradigm here is that it attempts to short-circuit the empirical process, appealing to hidden cognitive mechanisms instead of seeking useful perceptual variables that could underpin the perceptual achievement. A programmatic approach to direct social perception must be methodologically committed to the empirical strategy: if we have not yet identified the perceptual basis for the detection of some social property, we must assume that we have not looked hard enough (Warren 2021).

4 The Anscombean Strategy

One of the difficulties in studying psychology is that it requires great discipline on the part of researcher in keeping separate those things that we know and have access to, from those things that the actor we are studying knows and has access to. Confusing these two things seems obvious and easily avoidable, and yet it happens all the time. William James (1890) noticed this problem a long time ago. He called it the 'psychologist's fallacy'. Studying social cognition only compounds the difficulty here. In addition to the third-person perspective of the researcher and the first-person perspective of the actor, we are forced to add the second-person perspective of the social peer. As linguistically sophisticated adults we have a set of beliefs and desires that form an extensive 'theory' of how the world works. In social situations, we know how these come into play and are exhibited in our social behaviors. Because that is how our internal life works, we speculate that this is how others' minds work as well. When asked how we have the knowledge of others, we refer to this set of beliefs and desires we think they have. Thus, we explain social cognition in inferential terms.

The confusions are particularly acute in the domain of goal-directed behavior. In philosophy, goal-directed behavior (or, more broadly, intentionality) is mostly defined in terms of preceding mental states such as desires and beliefs (e.g., Bratman 1987; Davidson 2001). Since

attend to the movement that preceded the change. In the fifth and the sixth experiment, the researchers hypothesized that the stronger the animacy is indicated (by the degree of direction change or the speed of the movement), the faster the subjects will be to spot the change, which proved to be the case. These results suggest that animacy is distinguished in perception and prioritized in *bottom-up* allocation of attention during perception.

these approaches simply help themselves to speculating about hidden internal mechanisms, they are of little relevance to a direct social perception program. Fortunately, as argued by Segundo-Ortin and Kalis (2024), there is another way of thinking about goal-directedness that is potentially compatible with a direct social perception approach, namely Anscombe's theory of intentionality, which was inspired by the Wittgensteinian tradition of thought.

Segundo-Ortin and Kalis (2024) identify four prominent features of Anscombe's approach to intention. First, they point out that Anscombe takes a 'grammatical' approach towards intentionality and other problems of perception and action. That is, in the Oxford philosophical tradition, she attempted to explain perception and action by looking at the way they are reflected in language (Segundo-Ortin, Kalis 2024, 82). Although this is somewhat removed from the usual methods of ecological psychology, it may be useful for our programmatic aims. It could show us why other approaches (such as Davidson's or Bratman's) start from false intuitions about the phenomenon, and it could help us keep separate the different perspectives that are at play (first-, second-, third-person).

Second, and relatedly, Anscombe points out that actions are *only intentional under a description*. We can describe an action in many ways, all of which are valid. For example, when we are eating a cookie, we are decreasing the levels of cortisol in our blood, satisfying our hunger, stealing our colleague's dessert, increasing the demand for cookies in our country, etc. An action is intentional under a description if the action is identified as a means towards a goal according to this description. Intentional actions are, as quoted in Segundo-Ortin and Kalis (2024), "actions to which a certain sense of the question 'Why?' is given application; the sense is of course that in which the answer, if positive, gives a reason for acting" (Anscombe 2000, § 5, 9).

Third, intentional actions presuppose practical knowledge. Segundo-Ortin and Kalis (2024) describe three major characteristics of Anscome's understanding of practical knowledge. Firstly, we exhibit practical knowledge when we know what we are doing intentionally without observing it. Secondly, intentional action depends on the agent knowing what they are doing (possessing practical knowledge). Thirdly, practical knowledge allows us to answer 'why' questions about the performed action. As underscored in all three points, practical knowledge is inseparably connected to action. By conceiving goal-directedness in terms of practical knowledge, it is plausible that the goals in question could become accessible to a direct social perception approach.

Finally, intentional actions form a means-ends hierarchy or a teleological pattern. For instance, we might find ourselves at some point in the day filling a kettle with water, but we would not identify 'filling the kettle with water' as our goal. The action occurs as part of

a larger means-end hierarchy, in which the ultimate goal is making a cup of tea. Intuitively, actors are aware of this entire structure, in some sense: "by describing an action as intentional, what we are doing is indicating that the agent knows what she is doing right now, and how what she is doing embodies a certain teleological structure" (Segundo-Ortin, Kalis 2024, 80).

According to Segundo-Ortin and Kalis (2024), Anscombe's account provides a way of thinking about goal- directed action that avoids Cartesian dualism. They again quote Anscombe, clarifying the view of intentionality as a grammatical feature of actions: "an action is not called 'intentional' in virtue of any extra feature which exists when it is performed [...] [w]e do not add anything attaching to the action at the time it is done by describing it as intentional" (Anscombe 2000, § 19, 28). The relevant claim here, for our purposes, is that Anscombe's view avoids the assumption of invisibility. Intentional action is not a matter of any accompanying or preceding mental states of the agent but a realization of intention in action. In consequence, perceiving an intentional action can be direct and does not require 'peeking inside' another agent's mind.

A relevant empirical study here was conducted by Morris and Lewis (2010). This study looks at the phenomenon of diving (or simulation) in soccer, where players pretend to have been fouled by an opponent in order to gain an unfair advantage. The study shows that human observers are, at least in some circumstances, specialists in being able to detect an actor's real intentions. Participants were shown videos of soccer players falling to the ground following real collisions with opponents versus simulated collisions. The study found that the participants are not only in agreement about which ones of the presented videos showed real tackles, but that their judgment is mostly correct.[3]

Participants in this study can, in principle, notice features of the event such as a lack of temporal contiguity between the tackle and the fall, lack of ballistic continuity (e.g., does the falling player add motion to their roll on the ground that is not imparted by the tackle itself?), and lack of contact consistency (e.g., the player is struck in the chest and then falls down clutching their face). The ability of the participants to tell in which situations the player is pretending and in which their behavior is genuine lead the researchers to the conclusion that intentions can be perceptually detected. The study's results are consistent with the Anscombean view of intentionality. Because the players' intentions are realized in their actions, the players

3 The Morris and Lewis (2010) study addresses three questions: in the first part, the researchers try to find out whether participants agree about which players are pretending; in the second part, whether the participants judged correctly (their answers were compared to the instructions the players received before recording the dive on camera). In the third part, the researchers develop a taxonomy of diving behaviors.

cannot help but generate information about the deceit. This information is, both in principle and in practice, shown to be detectable by third-party observers.

The moral here is that in developing an empirical approach to direct social perception, it is imperative that we avoid conflating our description of the phenomenon of interest with the perspective of the actor or actors that we are investigating. Anscombe's approach is valuable because it provides a rigorous method for keeping the perspectives separate from one another. Anscombe shows that it is not necessary to understand intention in neo-Cartesian terms, with reference to hidden thoughts and desires. The observed agent, in their actions, provides us enough perceptual information so that we know their (realized in action) intentions. The Morris and Lewis (2010) study provides a useful example where the intention-in-action can be detected in information. This finding supports Anscombe's account: the intentions of the players seem to be indeed perceivable in their actions for the third-party observers. Therefore, there is no need to refer to internal mental states to explain intentions.

5 The Gestalt Strategy

A third positive strategy that has been advocated recently is to revive the practice of conceiving social perception in terms of Gestalts. This proposal has been raised in the domain of emotion perception by Forlè and Songhorian (2024) and by García Rodríguez (2021).

The paradigmatic example of a Gestalt is a melody, which consists of the arrangements of different notes but exceeds the sum of these notes played separately. A melody can be transposed to a different key such that none of the notes are the same as they were, and yet the melody itself remains intact. In Gestalt perception, we perceive the features of an object as a whole, not in isolation; the organization of the parts and the relations between them are crucial. Emotions, according to some direct social perception theorists, exhibit such a structure.

> [I]n cases of joy, one perceptually takes in a Gestalt that includes a particular kind of smile, wrinkled and glowing eyes, a generally relaxed bodily posture, certain linguistic expressions, and so on, in a context including relations to others and the environment; hence, one perceives a totality of features in context. (García Rodríguez 2021, 9440)

Expressions of emotions are both temporally and spatially extended, and exist across multiple sensory modalities. An emotion of joy might cause changes in posture, vocal pitch, or facial muscles. Gestalt

structures are actual perceptual phenomena, characterized by the following features: (1) they exceed the mere sum of their elements and (2) these elements gain new qualities in the context of the Gestalt (Forlè, Songhorian 2024).

The Gestalt view of emotions can be contrasted with the more standard cluster view of emotions. According to most contemporary theorists, emotions are clusters of, on the one hand, external properties such as facial expressions, and, on the other hand, internal properties such as feelings. As we saw above with Varga's (2020) thought experiment about the brain scanning machine for reading emotional states, this cluster view is compatible with the assumption of invisibility. The relevant parts of the emotional cluster may, after all, not be the parts that are visible. To truly perceive an emotion, on this view, we would have to perceive something that is private to the agent who is undergoing it. Therefore, direct perception of emotion is incoherent.

To escape this conclusion, Forlè and Songhorian (2024) suggest taking a different starting point. They suggest that a suitable view of interactions with others and our knowledge of their emotions can be found in the work of the German phenomenologist Max Scheler:

> According to Scheler, in our everyday encounters, we are neither confronted with a mere body nor with a mere mind, but with a psychophysical 'expressive unity' [...] This is the reason why Scheler can say that we are 'directly acquainted with another person's joy in his laughter, with his sorrow and pain in his tears, with his shame in his blushing, with his entreaty in his outstretched hands, with his love in his look of affection'. (Forlè, Songhorian 2024, 507)

It is possible that the argument here is overstating the extent to which the mainstream view of emotions is genuinely committed to dualism. Some direct social perception theorists have argued for a view of configural patterns that is potentially compatible with the cluster view. For instance, Krueger and Overgaard (2012) propose that certain bodily expressions are a *proper part* of the mental phenomenon, i.e. they are part of the emotional cluster.

That said, the strategy of stressing the Gestalt quality of emotions is, in our view, a useful one. Identifying the specific configural pattern that specifies a given emotional percept is exceedingly challenging (Zebrowitz, Collins 1997). Nevertheless, some such pattern must exist if a direct social perception program for emotional perception is to be viable. There is some inter-species evidence that suggests that such an approach is needed. Müller et al. (2015) report a study of dogs who are trained to respond preferentially to pictures of angry and happy human faces, which they are able to do at above chance levels. It is implausible that dogs could be using a theory of mind

strategy to achieve this. A logical place to look for the explanation would be in the dogs' sensitivity to higher-order Gestalt patterns.[4]

The moral that we would like to draw here is that a direct social perception program must address the issue of configural patterns. Social perception appears to be sensitive to higher-order configural patterns: in facial expressions, in bodily postures, in vocal modulation, etc. An adequate theoretical framework must be able to accommodate such configural features. A promising candidate framework already exists: the framework of Gibsonian information.

6 The Gibsonian Framework: Information for Direct Social Perception

As we have seen, recent debates around direct social perception have taken place largely among philosophers. Many discussions of direct social perception focus on a negative argument, namely, on whether we should reject representational interpretations of social perception in favor of a more embodied account (e.g., Gallagher 2020). We also saw that the issue of identifying the information for social perception is frequently conflated with the issue of describing interpersonal interaction. We then identified some positive strategies for developing a programmatic approach to direct social perception: the empirical strategy, the Anscombean strategy, and the Gestalt strategy. We suggest that empirical evidence supports findings in ecological psychology. Intentionality, understood in the Anscombean sense, points to a non-mentalistic picture of social cognition that avoids reference to hidden thoughts and desires and provides a view of intention-in-action detectable as ecological information. Emotions can be usefully understood as Gestalts and directly perceived as such, aligning with the principles of ecological psychology.

In concluding, we would like to suggest that all of these strategies can most naturally be brought together in the existing framework of James J. Gibson's ecological approach to perceptual information. We thus endorse the broad approach laid out in a target article by Becchio et al. (2018). The three pillars of empirical evidence, the Anscombean framework of intentionality, and the theory of Gestalts can serve as a starting point of the positive account of social ecological psychology.

Gibson is often invoked in discussions of direct social perception, particularly for his concept of affordances (e.g., Abramova, Slors

4 Müller et al. (2015) themselves posit that the mechanism is based on the dogs' memory for actual human faces, but this only postpones the explanation of how the dogs discriminated the meanings of those faces in the first place.

2015; Gallagher, Varga 2014; Kiverstein 2015; Valenti, Gold 1991). The concept of affordances is certainly relevant to any discussion of social cognition (Baggs 2021). However, our argument in the present paper is that there is a more fundamental vehicle for direct social perception that has been largely neglected in these recent debates, namely Gibsonian information (Becchio et al. 2018).

James J. Gibson wrote about the perception of social situations in the 1950s. Characteristically, Gibson sets out from the observation that perception works remarkably well most of the time. Gibson's argument here is that social perception research should aim to identify relevant properties of the stimulus that underlie this accurate perception:

> How do we perceive, for instance, that one person is being kind to another, bearing in mind that we do this with some accuracy? How do we perceive the intentions and abilities of a political candidate, taking it for granted that he does not fool all of us all the time? In other words, what do we discriminate and identify in these complex stimulus-situations which, when conditions are favorable, yields a correct perception? This ought to be the primary line of inquiry, but instead it is almost completely neglected. (Gibson 1951, 95–96, quoted in Heider 1958, 41)

This passage contains the kernel of a Gibsonian approach to social perception. The key to understanding Gibson's account of direct perception, generally, is that perception is a matter of discriminating structure that can be found within the stimulus-situation. Contrary to the main current of western philosophy, James and Eleanor Gibson argued that discriminating structure is sufficient for perception, and perception does not require mental augmentation in the form of, say, an act of mental categorization or the creation of a mental model (Gibson and Gibson 1955). Notice, once again, the similarity with Gallagher's (2008) glossing of direct perception as smart perception (see also Runeson 1977). Gibson later adjusted his terminology and argued for an account of perception as the direct detection of 'stimulus information' (Blau, Wagman 2022; Gibson 1979). In essence, the argument remains the same: perception research should be based on the detection or discrimination of informative structures within the stimulus-situation. Gibsonian information thus refers to discriminable structure present within the stimulus.

Gibson himself published little on social perception after the early 1950s (Reed 1988). Gibson's friend Gunnar Johansson, however, initiated an influential research program based on temporally-extended visual patterns that are generated by bodies in motion (Cutting, Kozlowski 1977; Johansson 1973; Runeson, Frykholm 1983). A Gibsonian approach to social perception must necessarily start from

such patterns within perceptual information (Baggs, Steffensen 2024; McArthur, Baron 1983).

We will not attempt here to present a complete account of how the Gibsonian perceptual information framework can be applied in the social context. We will offer only one illustrative example. Brett Fajen (2021), in a discussion of visual control of locomotion, raises the issue that when we move around in the world, we often encounter other objects that are themselves moving, and that we can potentially collide with.

> One factor that could complicate the perception of heading based on optic flow is the presence of objects that move independently of ourselves, such as pedestrians, automobiles, and cyclists. When such objects are present, as they often are, they introduce regions of the optic flow field with motion that is discrepant from that generated by the stationary background. In the instantaneous flow field, the velocity vectors corresponding to a moving object have different magnitudes and (in most cases) point in different directions. (Fajen 2021, 7)

Fajen asks us to imagine cycling into a traffic light-controlled road intersection as a vehicle crosses our path of locomotion, temporarily occluding the road ahead. As Fajen indicates in the passage above, we might expect that this situation would confuse our visual system. When cycling, we generate a centrifugal pattern of optic flow that has its centre in our current direction of heading. This information is held to be instrumental in controlling our action (Gibson 1958). The moving vehicle disrupts this pattern, temporarily denying us access to the useful source of information. Why do we not fall off the bike?

In practice, we are not confused by this situation. An object crossing our path is typically not experienced as a deletion of information, but as a presence of an object. The vehicle's movement is an additional source of information within the optic flow that exists alongside, or within, the self-generated global pattern of optic flow. The cyclist is well-advised to attend to this additional source of information. In the specific example that Fajen provides, the cyclist is on course to pass to the right of the vehicle, and this fact is specified in the leftward movement of the vehicle's solid angle as projected into the optic array of the cyclist. This complex configural pattern is information that the cyclist can use to detect that it is safe to maintain the current heading.

This example offers a useful illustration of how a Gibsonian approach to direct social perception might proceed. Most of us are familiar with the experience of moving around in traffic. Moving in traffic is a situation that, while it has its frustrations, we are nevertheless able to negotiate successfully most of the time. Given that

this is the case, we must be making use of a reliable source of information to guide our movements. The above discussion indicates that the information for social perception is, in fact, exceedingly rich. By paying appropriate attention, we can notice that other people and things are moving (we can detect animacy), and we can discriminate what goals are being pursued. We can also attend to features of the situation that are informative about other people's emotional states. In developing a programmatic approach to direct social perception, to paraphrase (Gibson 1951) information 'ought to be the primary line of inquiry.'

The direct social perception view today is primarily framed in negative terms, as a response to inferential, mentalist theories of mind. In this paper, we proposed three distinct strategies that support the framework of ecological psychology: empirical evidence, Anscombe's account of intentionality, and the Gestalt view of emotions. We suggest that these strategies position ecological psychology as a good candidate for a positive account of direct social perception. We admit, nonetheless, that the difficult work is yet to be done. The difficult work involves identifying and defining variables that specify social information in our environment. A goal for future work is to identify some concrete perceptual variables that might indicate the emotional states of others, and to investigate the extent to which such variables can be said to lawfully specify those emotional states. Ultimately, the viability of a Gibsonian approach to direct social perception hinges on whether such informational variables exist and are used by social animals.

Bibliography

Abramova, E.; Slors, M. (2015). "Social Cognition in Simple Action Coordination: A Case for Direct Perception". *Consciousness and Cognition*, 36, 519-31.

Anscombe, G. (2000). *Intention*. Harvard: Harvard University Press.

Baggs, E. (2021). "All Affordances Are Social: Foundations of a Gibsonian Social Ontology". *Ecological Psychology*, 33(3-4), 257-78.

Baggs, E.; Steffensen, S.V. (2024). "The Ecological Social Psychology of Aviation Disasters". *Ecological Psychology*, 1-17.

Becchio, C. et al. (2018). "Seeing Mental States: An Experimental Strategy for Measuring the Observability of Other Minds". *Physics of Life Reviews*, 24, 67–80.

Blau, J.J.; Wagman, J.B. (2022). *Introduction to Ecological Psychology: A Lawful Approach to Perceiving, Acting, and Cognizing*. New York: Routledge.

Bratman, M. (1987). *Intention, Plans, and Practical Reason*. Harvard: Harvard University Press.

Cutting, J.E.; Kozlowski, L.T. (1977). "Recognizing Friends by Their Walk: Gait Perception Without Familiarity Cues". *Bulletin of the Psychonomic Society*, 9, 353-6.

Davidson, D. (2001). *Essays on Actions and Events*. Oxford University Press.

De Jaegher, H. (2009). "Social Understanding Through Direct Perception? Yes, by Interacting". *Consciousness and Cognition*, 18(2), 535-42.

De Jaegher, H.; Di Paolo, E. (2007). "Participatory Sense-Making: An Enactive Approach to Social Cognition". *Phenomenology and the Cognitive Sciences*, 6, 485-507.

De Jaegher, H.; Di Paolo, E.; Gallagher, S. (2010). "Can Social Interaction Constitute Social Cognition?". *Trends in Cognitive Sciences*, 14(10), 441-7.

Dennett, D.C. (1978). "Beliefs About Beliefs". *Behavioral and Brain Sciences*, 1(4), 568-70.

Di Paolo, E.A.; Cuffari, E.C.; De Jaegher, H. (2018). *Linguistic Bodies: The Continuity Between Life and Language*. Cambridge, MA: MIT Press.

Fajen, B.R. (2021). *Visual Control of Locomotion.* Cambridge: Cambridge University Press.

Fernández Castro, V.; Heras-Escribano, M. (2020). "Social Cognition: A Normative Approach". *Acta Analytica*, 35(1), 75-100.

Forlè, F.; Songhorian, S. (2024). "Bodily Expressions as Gestalts. An Argument for Grounding Direct Perception Theories". *Philosophical Psychology*, 37(2), 505-27.

Gallagher, S. (2004). "Understanding Interpersonal Problems in Autism: Interaction Theory as an Alternative to Theory of Mind". *Philosophy, Psychiatry, & Psychology*, 11(3), 199-217.

Gallagher, S. (2008). "Direct Perception in the Intersubjective Context". *Consciousness and Cognition*, 17(2), 535-43.

Gallagher, S. (2020). *Action and Interaction*. Oxford University Press.

Gallagher, S.; Varga, S. (2014). "Social Constraints on the Direct Perception of Emotions and Intentions". *Topoi*, 33, 185-99.

Gallese, V.; Goldman, A. (1998). "Mirror Neurons and the Simulation Theory of Mind-Reading". *Trends in Cognitive Sciences*, 2(12), 493-501.

García Rodríguez, A. (2021). "How Emotions Are Perceived". *Synthese*, 199(3), 9433-61.

Gibson, J.J. (1951). "Theories of Perception". *Current Trends in Psychological Theory.* Pittsburgh: University of Pittsburgh Press, 85-110.

Gibson, J.J. (1958). "Visually Controlled Locomotion and Visual Orientation in Animals". *British Journal of Psychology*, 49(3), 182-94.

Gibson, J.J. (1979). *The Ecological Approach to Visual Perception*. Boston: Houghton Mifflin.

Gibson, J.J.; Gibson, E.J. (1955). "Perceptual Learning: Differentiation or Enrichment?". *Psychological Review*, 62(1), 32.

Gopnik, A.; Meltzoff, A.N.; Kuhl, P.K. (1999). *The Scientist in the Crib: Minds, Brains, and How Children Learn.* New York: William Morrow & Co.

Gordon, R.M. (1986). "Folk Psychology as Simulation". *Mind & Language*, 1(2), 158-71.

Heider, F. (1958). *The Psychology of Interpersonal Relations*. New York: John Wiley & Sons.

Heider, F.; Simmel, M. (1944). "An Experimental Study of Apparent Behavior". *The American Journal of Psychology*, 57(2), 243-59.

James, W. (1890). *The Principles of Psychology*. New York: Henry Holt and Company.

Johansson, G. (1973). "Visual Perception of Biological Motion and a Model for its Analysis". *Perception & Psychophysics*, 14, 201-11.

Kiverstein, J. (2015). "Empathy and the Responsiveness to Social Affordances". *Consciousness and Cognition*, 36, 532-42.

Krueger, J. (2012). "Seeing Mind in Action". *Phenomenology and the Cognitive Sciences*, 11, 149-73.

Krueger, J. (2018). "Direct Social Perception". *The Oxford Handbook of 4E Cognition*, 301-20.

Krueger, J.; Overgaard, S. (2012). "Seeing Subjectivity: Defending a Perceptual Account of Other Minds". *ProtoSociology: Consciousness and Subjectivity*, 47, 239-62.

McArthur, L.Z.; Baron, R.M. (1983). "Toward an Ecological Theory of Social Perception". *Psychological Review*, 90(3), 215.

Morris, P.H.; Lewis, D. (2010). "Tackling Diving: The Perception of Deceptive Intentions in Association Football (Soccer)". *Journal of Nonverbal Behavior*, 34, 1-13.

Müller, C.A. et al. (2015). "Dogs Can Discriminate Emotional Expressions of Human Faces". *Current Biology*, 25(5), 601-5.

Pratt, J. et al. (2010). "It's Alive! Animate Motion Captures Visual Attention". *Psychological Science*, 21(11), 1724-30.

Premack, D.; Woodruff, G. (1978). "Does the Chimpanzee Have a Theory of Mind?". *Behavioral and Brain Sciences*, 1(4), 515-26.

Reed, E.S. (1988). *James J. Gibson and the Psychology of Perception*. Yale: Yale University Press.

Richardson, M.J. et al. (2007). "Rocking Together: Dynamics of Intentional and Unintentional Interpersonal Coordination". *Human Movement Science*, 26(6), 867-91.

Rolfs, M.; Dambacher, M.; Cavanagh, P. (2013). "Visual Adaptation of the Perception of Causality". *Current Biology*, 23(3), 250-4.

Runeson, S. (1977). "On the Possibility of 'Smart' Perceptual Mechanisms". *Scandinavian Journal of Psychology*, 18(1), 172-9.

Runeson, S.; Frykholm, G. (1983). "Kinematic Specification of Dynamics as an Informational Basis for Person-and-Action Perception: Expectation, Gender Recognition, and Deceptive Intention". *Journal of Experimental Psychology: General*, 112(4), 585.

Scholl, B.J.; Gao, T. (2013). "Perceiving Animacy and Intentionality: Visual Processing or Higher-Level Judgment". *Social Perception: Detection and Interpretation of Animacy, Agency, and Intention*, 4629, 197-229.

Segundo-Ortin, M.; Kalis, A. (2024). "Intentions in Ecological Psychology: An Anscombean Proposal". *Review of Philosophy and Psychology*, 15(1), 69-89.

Spaulding, S. (2015). "On Direct Social Perception". *Consciousness and Cognition*, 36, 472-82.

Valenti, S.S.; Gold, J.M. (1991). "Social Affordances and Interaction I: Introduction". *Ecological Psychology*, 3(2), 77-98.

Varga, S. (2020). "Toward a Perceptual Account of Mindreading". *Philosophy and Phenomenological Research*, 100(2), 380-401.

Warren, W.H. (2021). "Information Is Where You Find It: Perception as an Ecologically Well-Posed Problem". *I-Perception,* 12(2), 2041-6695.

Westfall, M. (2023). "Perceiving Agency". *Mind & Language*, 38(3), 847-65.

Wimmer, H.; Perner, J. (1983). "Beliefs about Beliefs: Representation and Constraining Function of Wrong Beliefs in Young Children's Understanding of Deception". *Cognition*, 13(1), 103-28.

Zahavi, D. (2011). "Empathy and Direct Social Perception: A Phenomenological Proposal". *Review of Philosophy and Psychology*, 2(3), 541-58.

Zebrowitz, L.A.; Collins, M.A. (1997). "Accurate Social Perception at Zero Acquaintance: The Affordances of a Gibsonian Approach". *Personality and Social Psychology Review*, 1(3), 204-23.