On Whether We Can See Intentions

By Shannon Spaulding

(Forthcoming in Pacific Philosophical Quarterly)

Abstract: Direct Perception is the view that we can see others’ mental states, i.e., that we perceive others’ mental states with the same immediacy and directness that we perceive ordinary objects in the world. I evaluate Direct Perception by considering whether we can see intentions, a particularly promising candidate for Direct Perception. I argue that the view equivocates on the notion of intention. Disambiguating the Direct Perception claim reveals a troubling dilemma for the view: either it is banal or highly implausible.

1. Introduction

Can we see others’ mental states? We sometimes talk as if we can. We say things like, ‘I see what she is thinking’ or ‘I see his frustration.’ Is this just a way of talking, or can we really see others’ mental states? On the face of it, this is just a way of talking. We cannot see others’ mental states. We can see only their behavior and infer their mental states. Others’ behavior may be indicative – even strongly indicative – of their mental states, but we do not literally see their mental states, because mental states are not the kind of thing one literally can see.

Not everyone shares this perspective on the unobservability of mental states. Some theorists argue that sometimes we have direct access to others’ mental states. Theorists from embodied and extended cognition and cognitive neuroscientists argue that we can perceive others’ mental states with the same immediacy and directness that we perceive ordinary objects.

Consider what it is like to see an ordinary object in normal perceptual conditions. For example, I see that there is a coffee cup sitting on my desk. Seeing the coffee cup is immediate and direct. I do not have to infer from my sense data that there is a coffee cup on my desk. I
simply see it. Of course, seeing involves unconscious visual processing: light reflects off the object, my visual system detects surfaces and edges and constructs a 3-D representation of the object. However, there is an intuitive sense in which seeing the coffee cup is immediate and non-inferential. To grasp this, contrast seeing the coffee cup on the desk with seeing a Rorschach inkblot. Both involve unconscious visual processing, but the latter additionally involves interpretive or inferential mediation that the former lacks. It is in this sense that seeing the coffee cup is immediate and direct.

According to Direct Perception proponents, we can see mental states with the same immediacy and directness that we see ordinary objects, like the coffee cup on my desk. We do not need to infer from the observation of a target’s behavior that mental states (i.e., internal, unobservable causal states) are guiding the target’s behavior. In some cases, we simply see the mental states. I call this view Direct Perception.¹

My aim in this paper is to assess the plausibility of the idea that we can see mental states. However, there are several kinds of mental states – at a minimum, beliefs, desires, intentions, and emotions – and the plausibility of Direct Perception may differ depending on the kind of mental state discussed. To narrow the scope of the question of Direct Perception, I shall focus on whether we can directly perceive intentions. Intention is a particularly interesting mental state to discuss with respect to Direct Perception. Unlike most other mental states, intention has a conceptual connection to action. When I form an intention to do something, I commit myself to doing it. In forming the intention to work on this paper, I end the decision-making process about what to do and commit myself to working on this paper. Assuming no failures of memory, unforeseen obstacles, or new relevant information, I will work on this paper. There is a tight connection between particular intentions (the intention to Φ) and particular behaviors (Φ-ing).
I focus on intentions for three reasons. First, Direct Perception proponents explicitly claim that we can directly perceive others’ intentions. Second, the conceptual connection between an intention and its associated behavior makes intentions a promising candidate for Direct Perception. The idea that we can perceive a mental state by perceiving behavior requires, at a minimum, a tight connection between the mental state and the associated behavior. For most kinds of mental states, this requirement is not satisfied. Beliefs, desires, and complex emotions, such as guilt and shame, are not associated reliably with any particular behaviors.ii However, it is part of the concept of intention that an intention to Φ is correlated strongly with Φ-ing. Hence, intentions are a plausible candidate, perhaps the most plausible candidate, for Direct Perception. Third, if Direct Perception proponents fail to establish that we can directly perceive intentions, this result casts doubt on the more general Direct Perception account. If Direct Perception is not true even in the case of intention, then it is unlikely to be true with respect to other mental states, such as belief and desire, which are in no way reliably connected to particular behaviors. Thus, intention presents a good test case for Direct Perception.

2. Direct Perception

The Direct Perception accounts I consider come from three main camps: embodied cognition, extended cognition, and cognitive neuroscience.iii I shall briefly describe the basic argument for each account and detail the shared commitments of each of these accounts. Two points are worth making at the outset.

First, the views I canvass come from a diverse group of thinkers, and there is a danger that in lumping these views together the resulting account is one that no particular theorist holds.
I do not think that has happened here. In characterizing the view, I abstract away from the particular details and background theoretical commitments of each account and identify general principles shared amongst these various accounts. The general Direct Perception view that I describe is a simplified but fair representation of each account.

Second, in this section I shall make no evaluative judgments about the particular Direct Perception arguments I describe. Despite serious doubts about the details of many of these Direct Perception accounts, my aim is not to argue against them in a piecemeal fashion. I shall grant most of what they posit. In section 4, I argue that even if we grant the claims described in this section, the Direct Perception claim still is implausible.

2.1 Embodied Cognition

Embodied cognition is a relatively new research program in cognitive science that challenges cognitivism. Cognitivism has been the dominant view in psychology and philosophy of mind since the 1950s. Briefly, cognitivism holds that our cognitive capacities should be understood in terms of computational procedures operating on symbolic, internal mental states, and thus cognitive science should be focused on studying these internal states and processes. Embodied cognition rejects this account of the mind. According to embodied cognition, cognitivism allegedly makes the mistake of emphasizing the view of the mind as something to be studied independently of the body and its environment. Embodied cognition holds that emphasis in cognitive science should be on how the body and the environment shape the mind.

This view about how to study the mind has implications for social cognition. Embodied accounts of social cognition aim to explicate how our embodiment shapes our knowledge of others, and in what this knowledge of others consists. Shaun Gallagher and Daniel Hutto, two
prominent proponents of embodied social cognition, argue that intentions are not states of the mind, hidden away and inaccessible to observation. Rather, actions express intentions, and when I observe a target’s expressive behavior I observe his intentions.

In most intersubjective situations, that is, in situations of social interaction, we have a direct perceptual understanding of another person’s intentions because their intentions are explicitly expressed in their embodied actions and their expressive behaviors. This understanding does not require us to postulate or infer a belief or a desire hidden away in the other person’s mind. What we might reflectively or abstractly call their belief or desire is expressed directly in their actions and behaviors.

Gallagher and Hutto’s argument takes as its target two assumptions of traditional philosophy of mind: First, we start with the perception of mere movement. Second, in order to make the behavior meaningful, we theorize about the psychological states that could have caused it, or we draw an inference from ourselves to the other person in order to understand what the other is doing, feeling, or thinking. Gallagher and Hutto argue that both assumptions are mistaken. Others’ behaviors are not initially meaningless to us. They are imbued automatically with meaning in our social interactions. We do not need to make an inference and attribute a mental state in order to ascertain that social meaning. Intentions are expressed in embodied actions, and in perceiving others’ expressive embodied actions we directly perceive their intentions. For example, the intention to drink is expressed in grabbing a glass of water and bringing it to one’s lips. According to this view, in perceiving the drinking behavior, we literally perceive the intention to drink.
2.2 Extended Cognition

Extended cognition is the view that cognitive processes span the brain, body, and external world, and that cognitive states sometimes comprise parts of the external world. Some extended cognition theorists maintain that *intentions* span the brain, body, and external world. When we observe the external part of a target’s intention, his behavior, we perceive proper parts of his intention. Joel Krueger and Søren Overgaard argue:

Taking ‘expression’ in a constitutive sense is the idea that certain bodily actions are expressive of mental phenomena in that they actually make up proper parts of some mental phenomena. In other words, some mental phenomena have a hybrid structure: they straddle internal (i.e., neural) and external (i.e. extra-neural, gross bodily) processes. When we perceive certain forms of behavior and expressive actions, we quite literally perceive aspects of some mental phenomena.\(^{\text{vi}}\)

Krueger and Overgaard note that, ‘intentions are very often embodied in expressive actions, ripe for perception.’ In a more recent paper, Krueger further elaborates the Direct Perception claim. He explains, ‘[Direct Perception] claims that overt actions such as smiling, scowling, shaking one’s fists, gesturing while speaking, counting on one’s fingers, reaching for a beer, etc. give us direct perceptual access to other minds.’

The basic idea is that some mental states, including intention, have a hybrid internal/external structure. When we perceive the external parts of those mental states (the expressive behavior), we literally directly perceive constitutive, proper parts of the mental state.
2.3 Cognitive Neuroscience

In addition to arguments from embodied and extended cognition, theorists studying cognitive neuroscience argue that we can directly perceive others’ intentions. These theorists argue that mirror neurons are the mechanism of the direct perception of intentions.

Mirror neurons are found in the premotor cortex and the posterior parietal cortex, regions involved in sensory guidance of movement and the production of planned movement. A subject’s mirror neurons are involved in executing the subject’s own intentions. What has caused so much intrigue about mirror neurons, however, is that these neurons fire during the execution and observation of actions. In other words, a group of neurons (partially) responsible for producing and guiding planned movements also activates while one is observing others’ planned movements. It is argued that a subject’s mirror neuron activation has two roles: it realizes certain kinds of intention in the subject, and it is the mechanism for perceiving a target’s intentions. Thus, some theorists argue that mirror neurons are the neural basis for social cognition.

A number of theorists argue that mirror neurons are the mechanism of the direct perception of others’ intentions. Vittorio Gallese offers the most well developed account of how mirror neuron mechanisms underlie our perception of others’ mental states. According to that account, when a subject observes a target’s actions, the activation of the subject’s mirror neurons realizes a direct, experiential understanding of the target’s intentions. In Gallese’s terminology, mirror neuron activity realizes intentional attunement, a direct form of understanding others from within. Gallese argues that, ‘The activation of the mirror neuron system is intrinsically constitutive of action and intention understanding, at least at the level of basic actions.’

Theorists draw on a number of empirical studies as evidence for mirror neurons’ role in the direct perception of intentions. I shall describe just one of these experiments, as it is the most
frequently cited human mirror neuron experiment. The Tea Party experiment, as I will call it, is from a foundational study on motor mirror neurons in humans. Subjects observe 3-second videos of the following scenes: context, action, and intention. The context scene contains only objects, e.g., a teapot, a cup, a plate with cookies. The action scene portrays a hand grasping a cup without any contextual cues. The intention scene combines the context and action scenes, and subjects observe a hand grasping a cup in a scene with a teapot, a plate, cookies, etc.

The researchers found no mirror neuron activation while subjects observed the context scene, which is not surprising as there is no behavior to observe. They found some mirror neuron activation during the observation of the action scene, which is also not surprising, because we would expect mirror neurons to activate during the observation of a motoric behavior. The strongest activation occurred during the observation of the intention scene. This, they argue, is evidence that mirror neurons encode intentions.

The authors of the study hypothesize that if mirror neurons are sensitive only to the visual properties of the observed grasping action, activity in the mirror neuron areas should be equivalent while observing the action and intention scenes because the visual properties of the hand movement are the same in both scenes. If mirror neurons code the intentions associated with the observed action, thus predicting the actions that would follow the observed one, activity in the mirror neuron areas should be different between the action and intention scenes. And indeed there is higher activation for the intention scene than the action scene. The higher activation in the intention scene is due to the additional activation of so-called logically related mirror neurons, a subset of mirror neurons that fires in anticipation of the likely next behavior in a behavioral sequence. Thus, while observing the intention scene, logically related mirror neurons fire in anticipation of drinking. These theorists take the additional activation during the
intention scene as evidence that logically related mirror neurons code the intentions associated with observed actions.

In addition, half of the participants in this study were instructed to pay attention to the intention displayed by the behavior they were observing, while the other half were not told anything about intentions. The researchers found no difference in mirror neuron activation between the participants in each group, and in the debriefing session all participants were able to accurately report the intentions associated with each version of the intention scene. The authors of the study argue that this suggests that we associate intentions with actions automatically via the activation of mirror neurons.\textsuperscript{vii}

\textbf{2.4 Commitments of Direct Perception}

The accounts of Direct Perception I have described come from a heterogeneous group of thinkers who may have little else in common. Although the specific details and arguments for each account differ, the accounts converge on several commitments. In this section, I describe these shared commitments.

First, one important motivation for Direct Perception in all these cases is a rejection of a Cartesian-style asymmetry between understanding others’ mental states and understanding our own mental states. Descartes, among others, argues that we have immediate, privileged access to our own minds, but our access to other minds is indirect and epistemically less secure. Knowledge of other minds requires inferential mediation, which leaves open the possibility of massive error about other minds.

The Direct Perception argument against this Cartesian idea does not deny that there is some asymmetry between how we know our own minds and how we know other minds.\textsuperscript{viii} In
general, we have much more information about our own mental states than we do about others’,
we are more authoritative about our own mental states than about others’, and we access our own
mental states (but not others’) through introspection. Nevertheless, Direct Perception denies that
the first- and third-person cases are fundamentally different in a way that gives rise to the
Problem of Other Minds. It holds that we have direct access to our own mental states and to
others’ mental states, as well.\textsuperscript{ix} As a result, the Problem of Other Minds is shown to be a false
problem.

Second, these accounts hold that in ordinary social interactions our knowledge of others’
intentions is phenomenologically immediate. It seems that we simply see others’ intentions in
their behavior. A careful study of the phenomenology of social interactions reveals no experience
of inferring or calculating what another’s intention might be. We simply see your hand
movement as reaching for a piece of cake to eat, or your fumbling around in your purse as
searching for your keys.\textsuperscript{x}

Third, each of these accounts holds that we have non-inferential access to others’
intentions. The claim is that understanding a target’s intention does not require making an
analogical, causal, or abductive inference about the meaning of the behavior or what the target’s
intentions could be to have caused the observed behavior. Instead, we can know her intention
just by observing her behavior.\textsuperscript{xi}

Although Direct Perception theorists do not offer the following analogy, I think it
accurately captures what they have in mind. When I hear someone clearly and audibly speaking
English, I do not perceive the sounds and then make an inference from those sounds to the words
and sentences they express.\textsuperscript{xii} In normal cases, in perceiving the sounds, I directly perceive the
meaning of the sounds. The same allegedly applies to intentions. In ordinary cases of social
interaction, I do not perceive the behavior and make an inference to the intention the behavior expresses. In perceiving the behavior, I simply perceive the intention.

Direct Perception proponents do not provide an account of what an inference is, when an inference is required, and when it is not. They simply deny that in ordinary interactions knowing a target’s intention requires an inference from behavior to a mental state. Given the second and third commitments, that our knowledge of others’ intentions is phenomenologically immediate and non-inferential, one could be misled into thinking that Direct Perception proponents are assuming that all inferences are conscious. However, the second and third commitments are distinct. The second commitment concerns our conscious experience, and the third commitment concerns both our conscious and non-conscious experiences. To avoid this misunderstanding, I propose the following clarification of the third commitment: Understanding a target’s intention does not require a conscious or subconscious inference from behavior to mental state.

This does not exactly settle the matter of when an inference is required and when it is not. Dan Zahavi offers some help in clarifying when inferential mediation is not required. Some philosophers have reasoned that because social understanding depends on background context, social understanding is inferentially mediated. That is, in order to understand a target’s mental state, we have to take into account and make assumptions about background context. This may suggest that understanding others’ mental states involves inferring a mental state from the background information, context cues, and current behavior. Zahavi argues that this is a mistake. Direct Perception need not deny that knowledge of others’ mental states relies on contextual cues. He argues that it is perfectly consistent with Direct Perception to acknowledge that social understanding is influenced and enhanced by background knowledge, context cues, and past
experiences. Perception of an ordinary object, like the coffee cup on my desk, depends on background knowledge and context cues, but, it is argued, perception of ordinary objects is not inferential. Thus, Zahavi claims, inference is not required in virtue of the fact that understanding a target’s intention relies on background information and context cues.

Moreover, the fact that social understanding involves complex neural processing and multiple other cognitive processes (e.g., memory and perception) does not imply that social understanding requires inferential mediation. Perception of ordinary objects involves complex neural processing, but, according to Direct Perception, perceiving ordinary objects does not involve inference. Under normal conditions, we simply see the objects in the world. Language processing involves multiple other cognitive processes as well, but we do not have to infer what the sounds mean. We simply hear the sounds as meaningful language.

Social understanding involves multiple other cognitive processes, is neurally and computationally complex, and depends on background knowledge and contextual cues. All of this may be true, Zahavi argues, but none of it implies that social perception involves inference. This suggests that Direct Perception proponents are employing a narrow construal of what counts as an inference.

The fourth shared commitment of Direct Perception is that intentions have internal and external proper parts. This idea is an explicit commitment of extended cognition, and it is a tacit in the embodied cognition claim that behavior constitutively expresses intentions. Moreover, it is implied by the idea that mirror neurons realize intentions. Recall that the claim is that mirror neurons play two roles in this account: mirror neurons realize a subject’s intentions, and mirror neurons are the mechanism for directly perceiving a target’s intentions. It is the former claim that is relevant here. Mirror neurons are involved in the planning and production of actions. An
agent’s intentions are implemented through the activation of motor mirror neurons, which bring about behavior that is constitutive of the intention. The mirror neuron activation is the internal part, and the behavior that is realized from the mirror neuron activation is the external part. Hence, the hybrid character of intentions is implicit in the idea that mirror neurons realize intentions.

The fifth and final shared commitment is that mirror neurons are the mechanism of the direct perception of others’ intentions. This is the second (and more controversial) role for mirror neurons in Direct Perception. Both cognitive neuroscientists and embodied cognition theorists explicitly endorse this idea. Extended cognition theorists endorse this claim, as well. For example, Andy Clark claims that mirror neurons play an important role in understanding gestures. Gestures, he argues, are constitutive physical elements of a communicative cognitive process. Clark affirms the idea that we understand others’ gestures and, thus, the communicative intentions they express through the activation of mirror neurons.

Proponents of Direct Perception from various fields have converged on a similar cluster of claims regarding our access to other minds. Although it is plausible that adherents of each account endorse, either explicitly or implicitly, each commitment, we do not need such a strong claim. For our purpose, it is sufficient to show that proponents of these accounts have converged on a similar set of commitments. Below is a summary of this shared cluster of commitments. In the rest of the paper, I shall refer to the collection of these commitments as the Direct Perception account.

- Symmetry between first-person and third-person mindreading
- Phenomenological immediacy of intention understanding
• Non-inferential access to others’ intentions
• Intentions have internal and external proper parts
• Mirror neuron activation is the mechanism for the direct perception of intentions

3. Intentions

Direct Perception holds that we can see others’ mental states. I am considering whether we can see others’ intentions. Intention is a good test case for Direct Perception because it is one of the most plausible candidates for Direct Perception, several Direct Perception proponents explicitly say that we can directly perceive intentions, and if Direct Perception is not true even in the case of intention it is unlikely to be true in the case of other mental states, such as belief and desire, which are not reliably associated with particular behaviors. Thus, the question we face is whether we can see intentions.

Intentions come in three varieties: future-directed intentions, present intentions, and motor intentions. Direct Perception proponents do not specify exactly which kind of intention we allegedly are able to see. The evaluation of Direct Perception depends on which kind of intention is the target of Direct Perception, so in this section I shall offer a brief overview of the three varieties of intention.

A future-directed intention – e.g., the intention to call my dad on his birthday – is formed by making an explicit, conscious decision to perform action A. A future-directed intention is subject to certain rationality constraints. For example, the intention should be internally and externally consistent. That is, the various parts of the plan ought to be mutually consistent, and the intention must be compatible with what the agent believes about the world. Moreover, the
future-directed intention should be in accordance with the agent’s goals. The specific situational and temporal details of how one will carry out the future-directed intention are left indeterminate at this stage of the action plan. In forming the future-directed intention to call my dad on his birthday, my plan does not include information about whether I will be standing or sitting, calling from home or my office, calling in the morning or the evening, etc. I can have a future-directed intention to A without any outward behavioral signs indicating my intention to A. Formation of the future-directed intention terminates the decision-making process about what to do and initiates practical reasoning about the means of achieving A.

A present intention – e.g., the intention to call my dad here, now, on this phone – inherits the general action plan from the future-directed intention and makes that plan more specific. Although in both cases the content of the intention is intentional, the content of a future-directed intention is conceptual and descriptive whereas the content of a present intention is indexical and perceptual. Present intentions specify the means of achieving A, and the temporal and locational details of the action plan. Like future-directed intentions, present intentions are subject to conscious rational control. The agent exhibits rational control over present intentions by keeping track of one’s way of accomplishing an action and adjusting what one is doing to maximize one’s chances of success, and by controlling the side effects of accomplishing one’s action. If the side effects of accomplishing your action are negative, the rational thing to do is change the way of achieving your action or abandon the action altogether. The precise implementational details of the action plan are left indeterminate at this stage. For example, the plan does not specify whether I will dial my dad’s cellphone or home number, whether I will use iPhone’s Siri to dial or simply press the numbers manually, etc.
A motor intention – e.g., the intention to move my finger in such-and-such a way – implements the action plan inherited from the present intention.\textsuperscript{xviii} Motor intentions realize the intentional movement aimed at accomplishing the action. Motor intentions have lower-level guidance and control functions. That is, they have the function of (non-consciously) guiding and monitoring the action through inverse and forward models, and the function of controlling the implementation of the action, e.g., precision and smoothness of execution. Motor intentions are modular in that they are informationally encapsulated and domain specific. There is limited cognitive penetrability of motor intentions, and thus, unlike future-directed and present intentions, motor intentions are not subject to rationality constraints. The content of motor intentions is sensorimotor rather than conceptual and descriptive or indexical and perceptual.

The three kinds of intention differ in at least the following four ways: functional role, content, specification of action plan, and whether and how they are subject to rationality constraints. For a particular action, there may be temporal overlap amongst the kinds of intention, especially for present and motor intentions. In addition, not every action requires all three kinds of intention. Spontaneous actions admit of no real distinction between future-directed and present intentions, and habitual actions may require no future-directed or present intention at all. Nevertheless, these three kinds of intention play important roles in our actions.\textsuperscript{xix}

4. Dilemma

Now that we have all the relevant background information we can return to the issue of Direct Perception. Which kind of intention are we allegedly able to directly perceive? Direct Perception
proponents do not specify this, so we have to figure out which kind of intention fits best in the account.

It seems that motor intentions are the best candidate, for they uniquely satisfy the commitments of Direct Perception. Motor intentions have a hybrid structure that straddles internal (neural) and external (extra-neural, gross bodily) processes. When one has a motor intention, internal neural processing realizes external motor behavior. The appeals to mirror neurons as the mechanism of the direct perception of intention further indicates motor intention, as mirror neurons activate in response to low-level aspects of motoric behaviors. If mirror neurons are the mechanism of the direct perception of motor intentions, this makes sense of the phenomenal immediacy claim. Mirror neuron activation is automatic; it requires no conscious, phenomenologically salient deliberation about the goal directedness of the behavior. Moreover, this suggests that perceiving the external expression of others’ motor intentions is not inferentially mediated. We simply see other’s movements as goal-directed, e.g., as grabbing, pushing, or pulling.

If there is a form of intention that satisfies the commitments of Direct Perception, it is motor intentions. The idea is that by perceiving a proper part of a motor intention (the externally-expressed behavior), we literally directly perceive the motor intention. For example, in perceiving your hand movements, I directly perceive your motor intention to grasp an object. Thus, I argue that the most charitable interpretation of the idea that we can see intentions is that we can see motor intentions.

Present intentions seem to be the wrong target for Direct Perception, for they do not satisfy the shared commitments described above. Present intentions may be phenomenologically immediate, but they need not be. Similarly, present intentions may be accompanied by externally
observable behavior aimed at realizing those intentions, but they need not be. My behavior may not make evident that I have the present intention to call my dad right now. If am prevented from enacting this intention by external obstacles (e.g., my phone is missing) or internal obstacles (e.g., I get distracted and forget what I was doing, or I have an akratic disposition), there will be no external behavior realizing my present intention to call my dad. I can have the sincere present intention to call my dad now without engaging in any externally observable behavior, which indicates that present intentions do not have external proper parts. Moreover, access to a target’s present intentions is inferentially mediated. Suppose you know that today is my dad’s birthday, that I am pretty reliable about calling my family members on their birthdays, and you see me reach for my phone. It is not obvious from the background context and my behavior that I intend to call my dad right now. There are a number of other plausible present intentions that I may be acting on. Knowing that I am acting on a particular present intention, to call my dad, requires an inference. Finally, it is not clear how mirror neurons could be the mechanism for perceiving present intentions. This would require an as-yet undeveloped account of how mirror neurons detect or are somehow sensitive to mental states that need not be associated with any outward behavior. We know how mirror neurons detect motor intentions. Based on the kinematics of, say, a hand movement (i.e., the position of the hand relative to an object, the angle and speed of the hand movement, and the sensory properties of the object), mirror neurons activate in anticipation of the most likely next motor movement in the behavioral sequence. It is hard to see how a story like this could be told for present intentions, such as the intention to call my dad. Thus, present intentions are the wrong target for Direct Perception because they do not satisfy the commitments of the view.
Future-directed intentions are a poor candidate for Direct Perception, as well. It is not at all clear how one literally could observe a future-directed intention. Future-direct intentions – explicit, conscious, rationally constrained decisions to perform some action – are not the sort of thing one can perceive. A target may announce his future-directed intention, but surely that is not what Direct Perception theorists have in mind. One may infer that a target has a particular future-directed intention, but one cannot literally perceive the future-directed intention. Furthermore, future-directed intentions do not satisfy the commitments of Direct Perception. Access to others’ future-directed intentions often is not phenomenologically immediate. Our knowledge of others’ future-directed intentions always is inferentially mediated. Future-directed intentions do not have external proper parts, and clearly mirror neurons could not be the mechanism of the direct perception of future-directed intentions.

Given that motor intentions obviously satisfy the commitments of Direct Perception whereas present and future-directed intentions do not, motor intentions seem like the best candidate for Direct Perception. We directly perceive the constitutive behavioral expression of a motor intention. But if this is what Direct Perception amounts to, the view is not that radical. Our perception of the expression of others’ motor intentions is phenomenologically immediate and non-inferentially mediated. We automatically perceive some behavior as, for example, grasping, pushing, or pulling. This idea is compatible with the claim that much of our ordinary social understanding involves inferring others’ present and future-directed intentions.

Direct Perception proponents face the following dilemma. Either the target of Direct Perception is motor intentions, or it is some other kind of intention. On the one hand, if the target is motor intentions, then the view is not very controversial. Essentially, the claim that we directly perceive motor intentions amounts to the idea that from the perception of current motor
movements, we anticipate the next motor movement in a sequence. In neural mechanistic terms, based on the kinematics of the hand movement and sensory properties of a nearby, observable, graspable object, our brains anticipatorily activate as if we are perceiving a target grasp an object. This neural mechanism makes possible the direct perception of motor intentions. In this way, we immediately, non-inferentially perceive a target’s hand movement as a grasping.

I do not claim that everyone accepts this account of how we perceive motoric behavior, but it is not a controversial view. It does not depart significantly from orthodox views in philosophy or theoretical psychology. Most troubling for Direct Perception proponents, it is compatible with accounts that endorse a logically and epistemically important asymmetry between self-knowledge and knowledge of other minds. In fact, a modern-day Cartesian could perfectly consistently accept this idea. Moreover, the idea is compatible with the view that access to our own minds and other minds is indirect and inferentially mediated. Thus, although the phrasing of the idea is unusual, the direct perception of motor intentions hardly is a radical departure from the view that our access to other minds is indirect and inferentially mediated.\textsuperscript{xxi}

On the other hand, if Direct Perception is committed to the idea that we can directly perceive present and/or future-directed intentions, then the view is highly implausible. Present and future-directed intentions do not satisfy the commitments of Direct Perception, and we have no good reason to think that Direct Perception even could apply to present and future-directed intentions. If Direct Perception proponents take this second horn of the dilemma, at the very least we are owed a story about how exactly it is possible to directly perceive present or future-directed intentions.

5. Epistemology of Perception
In the previous section, I argued that we do not directly perceive future-directed and present intentions in just the same way that we directly perceive ordinary objects in the world. At best, we directly perceive the external expression of motor intentions. This result is bound to be dissatisfying for Direct Perception proponents, who aim to challenge the standard view of how we know others’ intentions. In this section, I consider whether the notion of basic perceptual belief from the epistemology of perception literature can vindicate Direct Perception.

Basic beliefs are epistemologically basic in that they are non-inferentially justified. Basic beliefs contrast with non-basic beliefs, which require justification via inference from other beliefs. Perceptual beliefs, some argue, are a species of basic beliefs. Believing that you are seeing the color red is a basic perceptual belief. To be justified, this belief does not require an inference from other basic or non-basic beliefs. Perceptual beliefs about ordinary objects may be basic beliefs. The perceptual belief that there is a coffee cup on my desk may be a basic perceptual belief. Perhaps, a Direct Perception theorist could argue, the belief I acquire about your intention is a basic perceptual belief. That is, I observe your behavior and in doing so acquire the belief that you intend to \( \phi \). One could argue that this belief is a perceptual belief that is non-inferentially justified. My perceptual belief about your intention requires no inference from other beliefs about your behavior, the environment, background context, etc. I just see you as intending to \( \phi \). If this were the case, it would vindicate the idea that we perceive others’ intentions with the same immediacy and directness that we perceive ordinary things in the world.

The idea that beliefs about others’ intentions are basic perceptual beliefs differs from the arguments that Direct Perception proponents offer. It is, however, a possible way to resuscitate the idea that we have direct, immediate access to others’ intentions. If we have basic perceptual
beliefs about others’ intentions, then these beliefs are produced automatically and non-inferentially. I shall argue that this maneuver will not help the Direct Perception theorist. In fact, consideration of the idea of basic perceptual beliefs reinforces my argument from section 4 that Direct Perception is either banal or highly implausible.

What exactly is a basic perceptual belief? According to one account, what distinguishes basic beliefs from non-basic beliefs (i.e., inferential cognition) is the nature of the cognitive system that produces the beliefs. For basic perceptual beliefs, the distinguishing factor is the nature of the perceptual system that produces the perceptual beliefs. Jack Lyons argues that in order to be a perceptual module or system capable of producing basic perceptual beliefs, a cognitive system must satisfy the following four criteria:

a. Its lowest-level inputs are transductions across sense organs.

b. None of its inputs to any of its subsystems is under direct, voluntary control of the larger organism.

c. It is inferentially opaque, i.e., none of its inter-level representations are conscious beliefs.

d. It has a ‘normal’ etiology, i.e., it results from the interplay of learning and innate constraints.

A system that satisfies criteria (c) and (d) is what Lyons calls a primal system. Primal systems admit of inferential and non-inferential operation. A primal system operates inferentially if it takes beliefs as input and bases its output on these beliefs. In contrast, a primal system operates non-inferentially if it takes something other than beliefs (e.g., queries, conceptions, etc.) as input.
We now have the resources to define basic perceptual belief. A perceptual belief is basic if and only if it results from a cognitive system that satisfies the above four criteria and does not take beliefs as input.

In section 2, I discussed Zahavi’s comments about when inferential cognition is not required. Zahavi’s comments were helpful, but they did not tell us when inferential cognition is required. Lyons’ account fills this gap. It provides criteria for determining when cognition is inferentially mediated. A cognitive process is inferentially mediated when it fails to meet one or more of the above criteria.

A belief about another’s intention is a basic perceptual belief if and only if it is produced by a cognitive system that satisfies the above requirements. Could beliefs about others’ intentions be produced by such a system? Again we have to ask which kind of intention. Beliefs about others’ future-directed intentions surely are ineligible. These beliefs are not produced by a perceptual system with features a – d. First, there are multiple cognitive systems involved in producing a belief about a target’s future-directed intention, not all of which are perceptual systems. Second, the cognitive process that produces beliefs about future-directed intentions involves conscious deliberation, is flexible, and is under the direct, voluntary control of the cognitive agent. I can choose to deliberate on someone’s future-directed intention. For example, when a ‘frenemy’ does something nice for me, I can choose whether or not to try to figure out his future-directed intention. When I am feeling charitable, I may decide to take the nice deed at face value. When I am feeling cynical, I may try to discern my frenemy’s intentions. Moreover, in coming to a belief about my frenemy’s future-directed intention, I take as input my conscious beliefs about his personality, his past behavior, and his goals. Thus, for several reasons, beliefs about others’ future-directed intentions are not basic perceptual beliefs.
Beliefs about others’ present intentions also are unlikely candidates for basic perceptual beliefs. To be basic perceptual beliefs, the beliefs must be produced by a cognitive system whose operations are automatic, not subject to voluntary control, not consciously accessible, and do not take beliefs as input. This simply is not the case for the cognitive process that produces beliefs about others’ present intentions. We can wonder what someone is doing and consciously think about what present intention her behavior is serving. For example, as you are rummaging through your purse, I may or may not wonder what you are trying to do. Coming to a belief about your present intention is not a mandatory, automatic cognitive process. Rather, it is under the voluntary control of the agent. If I am motivated to figure out what you are trying to do, I may notice that the server has just handed you a check for your meal and infer that you are trying to find your wallet. As the previous example illustrates, the cognitive process is consciously accessible. When the agent is motivated to figure out the target’s present intention, and when the present intention is not obvious, the process rises to the level conscious deliberation. The cognitive process that produces beliefs about others’ present intentions is consciously accessible, non-mandatory, and can take beliefs as input, which indicates that it does not satisfy Lyon’s criteria. Thus, beliefs about others’ present intentions are not basic perceptual beliefs.

Again, motor intentions seem like the most plausible candidate. The mirror neuron system, the cognitive system that produces perceptual beliefs about (the external expression of) others’ motor intentions, satisfies the above criteria. It takes as input transductions across sense organs, not consciously accessible beliefs. The operation of the mirror neuron system is informationally encapsulated, automatic, and mandatory, thus not under direct, voluntary control of the agent. Moreover, the inter-level representations are sensorimotor representations, not consciously accessible beliefs. The mirror neuron system is a very good example of a perceptual
primal system that operates non-inferentially. If we can have basic perceptual beliefs about another’s intention, the only plausible candidate is basic perceptual beliefs about motor intentions (more precisely, basic perceptual beliefs about the goal-oriented behavior that is the expression of the motor intention).

Perceptual beliefs about goal-directed behavior that are produced (in part) by the mirror neuron system satisfy the criteria for basic perceptual beliefs, whereas perceptual beliefs about present and future-directed intentions do not. This result simply reinforces the dilemma described in the previous section. Either the target of Direct Perception is motor intentions, in which case the view is not a radical alternative to traditional views that it is meant to challenge. Or, the target of Direct Perception is present or future-directed intentions, in which case the view is highly implausible.

6. Conclusion

The idea I consider in this paper is that we can perceive others’ intentions with the same directness and immediacy that we perceive ordinary objects in the world. I have argued that all of these arguments for Direct Perception have the same basic problem: they equivocate on the notion of intention. How we evaluate Direct Perception depends on what kind of intention we allegedly are able to directly perceive. The only kind of intention that is remotely plausible and satisfies the commitments of Direct Perception is motor intentions, but the direct perception of motor intentions is not a serious departure from the view that mental states are inferred, not observed.
Thus far, Direct Perception proponents have not distinguished the different kinds of intention and have not specified the precise target of Direct Perception. The upshot of this discussion is that specifying the target of Direct Perception is essential for evaluating the view. The conclusion of the Direct Perception argument either is the fairly uncontroversial idea that we can see the externally realized parts of others’ motor intentions, or it is the implausible idea that we literally can see others’ present or future-directed intentions. Furthermore, the failure to establish that we can directly perceive intentions in any interesting sense spells trouble for the general Direct Perception account. Given that Direct Perception fails in one of best-case scenarios, it is unlikely to succeed for other kinds of mental states, such as beliefs and desires. Thus, this result is a serious blow to the Direct Perception account.

Department of Philosophy
Oklahoma State University

REFERENCES


NOTES

---

i Direct Perception is not behaviorism. The argument is not that all mental states are identical or reducible to behavior. Direct Perception proponents argue that some mental states are observable. This is compatible with the idea that certain types of mental states may be entirely covert, while others are more or less observable. The Direct Perception claim also is not an instance of the familiar argument in philosophy of science that, after sufficient training, one can observe theoretical objects, such as electrons or quarks. Direct Perception proponents do not deny that observation is theory laden or that we can come to observe theoretical objects or events. They deny that this is what is happening in direct perception of mental states.

ii It is plausible that some basic emotions, such as fear, anger and disgust, reliably are associated with particular physiological responses and behaviors. Thus, these basic emotions also may be a good test case for Direct Perception. In another paper (Spaulding (in progress)), I argue that current discussions of Direct Perception are hindered in lacking an account of inference and ambiguity about the positive view of Direct Perception. It is unclear whether the view is that the content of perception includes representations of mental states or that the content of basic perceptual beliefs includes representations of mental states. I argue that it is plausible that we directly perceive basic emotions in the latter sense.

iii Other philosophers argue for a view similar to this, but their arguments really are quite different from the arguments considered here. See, for example, McDowell (2004), McNeill (2012), Smith (2010), and Strawson (1959). In order to address a unified Direct Perception account, I shall not consider these views in this paper.

iv Although, for persuasive criticisms of various elements of Direct Perception, see Herschbach (2008), Jacob (2011), and Lavelle (2012).


vi Joel Smith (2010) offers a somewhat similar view. He argues that what is presented to the observer is the target’s behavior, but there is more to our experience than just what is visually
presented. What you visually experience is not exhaustive of what you see. The mentality of a
target’s behavior is co-presented with the behavior. What is presented and co-presented take part
in a harmonious experience, the result of which is that we see both at the same time.

vii For a critique of this interpretation of mirror neurons and an alternative account of their
role social cognition, see Spaulding (2013) and Jacob (2008), Jacob (2013).

viii Unlike the views discussed in the main text, Max Scheler Scheler (1913/2007), an
early proponent of Direct Perception, holds that there is no asymmetry between our access to our
own and others’ mental states. On his view, we perceive our own mental states in our actions and
expressive behaviors in the same way that we perceive others’ mental states in their actions and
expressive behaviors. Thanks to an anonymous referee for highlighting this historical fact.

ix There are, of course, more than two options here. The Cartesian view holds that we
have direct access only to our own minds, whereas the Direct Perception view holds that we have
direct access to other minds. Following Peter Carruthers (2011), one could argue that we do not
have direct access to either our own minds or other minds. On this view, which I find quite
plausible, knowledge of our own minds is just as inferential and interpretive as knowledge of
other minds.

x You do not need to subscribe to Direct Perception to think that this is the right analysis
of our phenomenology. Opponents of Direct Perception can, and probably do, endorse this claim
about our phenomenology in ordinary social interactions.

xi This claim does not imply that others’ behavior always is immediately explicable to us.
Sometimes we have no idea what someone is trying to do, and we have to come up with possible
explanations of the behavior. But, according these accounts, such cases are relatively rare.

xii Language processing involves unconscious processing that some people regard as
inferential. Recall from section 1 my discussion of seeing the coffee cup on my desk. Seeing the
coffee cup involves unconscious processing, but not all unconscious processing is inferential in the
relevant sense. I shall discuss this point later in this section and in section 5.

xiii For example, Krueger and Overgaard (2012) are careful to distinguish conscious from
non-conscious inferences.

xiv On some accounts, mirror neurons are agent neutral, i.e., they encode the same
information about a behavior whether it is the subject’s behavior (in execution mode) or the
target’s behavior (in observation mode). With respect to Direct Perception, this entails that the
inner part of the hybrid intention is not self-specific. This may be a different sort of hybridity
than, say, emotions where the inner (physiological) part is self-specific. However, whether or not
mirror neurons are agent neutral is a controversial matter. There are in fact differences in mirror
neuron activation (and the larger mirror neuron system) in observation mode and execution mode
Spaulding (2012). Thus, what sort of hybridity is involved here is an open question. Thanks to an
anonymous referee for bringing to my attention this possible complication.

xv Unlike the other commitments, the claim that mirror neurons are the mechanism of
direct perception is not a conceptual commitment. It is an open empirical possibility that some
other neural mechanism is a better candidate for Direct Perception. Thus far, most Direct
Perception theorists have pinned mirror neurons as the most plausible neural mechanism, but this
claim is subject to empirical evaluation.

xvi I do not endorse these commitments. Nevertheless, I shall not challenge these
individual claims here. In sections 4 and 5, I offer a more fundamental critique of Direct
Perception.
Most theories of action posit future-directed and present intentions. I follow Pacherie in advancing an account with future-direct, present, and motor intentions. As Pacherie notes, properties of motor intentions are often attributed to present intentions, which generates confusion about whether present intentions are under rational, voluntary control. Distinguishing between present and motor intentions alleviates this confusion. More importantly for our purpose, distinguishing between present and motor intentions brings into focus the exact commitments of Direct Perception.

Jeannerod (1994) is a foundational empirical article on motor representations (what I call motor intentions) and their relationship to semantic representations (including present and future-directed intentions). For a review of the recent empirical findings on motor intentions and how they relate to other sorts of intentions, see Hamilton and Grafton (2007) and Rizzolatti and Sinigaglia (2010).

For more on these three kinds of intention, see Pacherie (2006). In a recent article, Butterfill and Sinigaglia (2014) argue that motor representations are not a kind of intention. Despite appearances to the contrary, their view differs only terminologically from the view advanced here. They argue that intentions have a propositional representational format whereas motor representations do not. This implies that motor intentions cannot serve as a premise or conclusion in practical reasoning, and thus they have a different functional role than (future-directed and present) intentions. Nevertheless, they argue, the outcomes specified by intentions and motor representations are closely related, and a complete account of action must show how intentions are related to motor representations. I agree with all of this. If Butterfill and Sinigaglia are right that we should not call these motor representations motor intentions, then my conclusion is that the best that Direct Perception can hope for is the claim that we have direct access to external expressions of others’ motor representations. Similar remarks apply to motor representations in Jeannerod’s sense Jeannerod (1994). One could substitute motor representation for motor intention in my argument, and the conclusion would remain intact: Direct Perception either plausibly holds that we non-inferentially perceive external proper parts of motor representations, or it implausibly holds that we non-inferentially perceive present or future-directed intentions (semantic representations, in Jeannerod’s terminology).

Corroborating this interpretation, Rizzolatti and Sinigaglia state, ‘The studies reviewed above indicate that the parieto-frontal mirror network may subserve the understanding of the motor intention underlying the actions of others… This does not mean that the parieto-frontal mirror mechanism mediates all varieties of intention understanding. Intention understanding is a multi-layer process involving different levels of action representation, from the motor intention that drives a given chain of motor acts to the propositional attitudes (beliefs, desires and so on) that — at least in humans — can be assumed to explain the observed behaviour in terms of its plausible psychological reasons Rizzolatti and Sinigaglia (2010).’

If one is skeptical of motor intentions, one could substitute motor representations Butterfill and Sinigaglia (2014), Jeannerod (1994). See endnote 19. The dilemma would remain for Direct Perception, though. Either we directly perceive external expression of motor representations, or we directly perceive present or future-directed intentions. The former is relatively banal, and the latter is implausible.

This is meant to account for basic beliefs from introspection, memory, and a priori intuition, in addition to perception.
This is an oversimplification. The mirror neuron system receives input from various cognitive systems (the visual system) and operates in conjunction with other systems and neural regions (e.g., the superior temporal sulcus) to produce beliefs about motor intentions. None of these other systems violates Lyon’s criteria, so for our purposes here this is a harmless oversimplification.

I am grateful to Robert Lurz, Karen Neander, Katherine Rickus, Sarah Robins, Michael Roche, and Michael Wilby for detailed feedback on various drafts of this paper. Thanks to the audiences at Texas Tech University, University of Memphis, the Metaphysics of Mind and Brain workshop at Humboldt University in Berlin, and the Southern Society for Philosophy and Psychology in Austin, TX. The questions from the audience helped me clarify the target and structure of my argument. Finally, thanks especially to the two anonymous reviewers at this journal for their helpful comments, critiques, and suggestions.