ANDREA BLOMQVIST
*University of Sheffield*
*ablomqvist1@sheffield.ac.uk*

# DIRECT SOCIAL PERCEPTION OF EMOTIONS IN CLOSE RELATIONS[1]

*abstract*

*Drawing on a pluralist approach to mindreading, I explore Direct Social Perception with respect to perceiving the emotional states of people that we are close to, such as spouses, friends, and family. I argue that in general, emotions are embodied and can be perceived directly. I further claim that perceptual content includes concepts. That is, I argue against a non-conceptual view of emotion recognition, claiming instead that we learn emotional concepts by attending to certain expressive patterns of emotions. This view implicates that we can directly perceive both basic and non-basic emotions of people we are close to.*

*keywords*

*direct social perception, embodied emotions, mindreading, pattern recognition*

Phenomenology and Mind, n. 12 - 2017, pp. 184-195
DOI: 10.13128/Phe_Mi-21117
Web: www.fupress.net/index.php/pam


ISSN 2280-7853 (print) - ISSN 2239-4028 (on line)

Direct Social Perception (DSP) is the idea that we can directly perceive others' mental states. Proponents of DSP claim that some mental states – usually motor intentions and emotions, but also sometimes beliefs (Gallagher and Hutto, 2008) – are embodied, and can therefore be directly accessed by others using perception (Gallagher, 2008; Gallagher, 2015; Krueger and Overgaard, 2012; Newen, Welpinghus, and Juckel, 2015; Spaulding, 2015; Zahavi, 2011). This paper will exclusively deal with emotions – in particular the non-basic emotions[2] of people that we know well.

DSP is often contrasted with inferentialist mindreading theories such as Theory Theory (TT) and Simulation Theory (ST).[3] Both of these theories claim that others' mental states are hidden from us; we need to use a theory or simulation to gain access to others' mental states, and we can only non-inferentially gain access to our own mental states. According to TT, we do this by using folk psychological rules such as "if A wants p and believes that doing q will bring about p, then ceteris paribus, A will do q" (Borg, 2007). When observing someone reaching for a bottle, I attribute the desire that they want the bottle to them, and the belief that reaching for it is the best way of achieving their goal. Similarly, if I observe someone smiling as they greet a friend, I may attribute happiness to them. Thus, by observing behavior, I infer that an agent has certain mental states. According to ST, we understand others' mental states by simulating their emotions or thoughts offline in ourselves and then projecting the simulated state to them (Goldman, 2006). That is, we put ourselves "in their shoes" to understand them.

In this paper, I will argue for a pluralist approach to mindreading.[4] That is, rather than claiming that the above theories are incompatible, I will claim that we use different strategies in different contexts depending on which is the least effortful to use (Fiebich and Coltheart,

1 I am grateful to Will Hornett who on numerous occasions gave me insightful feedback on my manuscript as well as took time to discuss it with me – this significantly improved the paper. I am also grateful to my supervisor Luca Barlassina and the two anonymous referees who reviewed the manuscript and provided useful comments.

2 I use 'non-basic' and 'complex' emotions interchangeably, where I take them to be any emotion that is not basic. Basic emotions include happiness, surprise, anger, sadness, fear, and disgust (Ekman, Friesen, and Ellsworth, 1972). Examples of non-basic/complex emotions are jealousy, sorrow, pride, etc.

3 See for example Gopnik and Wellman (1992) for a defense of a version of TT, and Goldman (2006) for a defense of a version of ST.

4 For pluralist accounts of mindreading, see Fiebich and Coltheart (2015); Gallagher (2015); and Fiebich, Gallagher, and Hutto (2016).

2015, p. 249). I will focus on the context[5] of interaction in close relations and argue that we use DSP in face-to-face interactions with people that we know well (e.g. spouses, friends, family).[6] My account will thus focus on interactions between people who are well acquainted, something which is not traditionally discussed by TT and ST (Fiebich and Coltheart, 2015, p. 236). There is a strong reason for the mindreading literature to focus on these cases, since we actually spend a significant amount of time mindreading people that we are close to, rather than strangers (think for example about how frequently spouses interact and how mindreading is important for the relationship to work smoothly). In these kinds of close relations, I will further claim, we can not only come to perceive the basic emotions, but we also perceive non-basic (sometimes called "dispositional") emotions, such as jealousy and pride.[7] I will argue that this is a process that is based on pattern recognition as suggested by Newen, Welpinghus, and Juckel (2015). However, I will depart from their view of the process as non-conceptual (Newen and Schlicht, 2009), and instead argue for a conceptual account based on Siegel's theory of rich perception and Wu's theory of attention (Siegel, 2006; Wu, 2008). This paper will thus develop an account of how we gain the ability to directly perceive non-basic emotional states of people we are close to.

In Section 1, I will outline the basic claims of DSP, considering how we can come to perceive emotional states by perceiving embodied emotions. Following Green, I will claim that we can perceive a mental state by perceiving part of it (Green, 2010). In Section 2, I will argue that perception is rich, and that we do not only perceive low-level properties like shapes and colors, but we can also perceive high-level states such as emotions, provided that we possess the right emotional concepts (Siegel, 2006; 2010). To explain how these concepts are acquired, I will draw on Wu's account of attention (Wu, 2008). In Section 3, I will consider Newen and Schlicht's argument that emotion recognition is based on non-conceptual pattern recognition (Newen and Schlicht, 2009). Here, I will be sympathetic to the idea that pattern recognition is necessary for emotion recognition, but argue that this process is conceptual. I will back this up by empirical data showing that people with dementia have impaired emotion recognition because they lack emotional concepts. Finally, in Section 4, I will show how rich perception, attention, and pattern recognition can enable us to directly perceive the non-basic emotions of people that we are close to.

## 1. Directly Perceiving Embodied Emotions

DSP theorists claim that mindreading is perceptual rather than cognitive and that we can directly perceive certain mental states of other people. DSP, they argue, is a non-inferential kind of mindreading that stands in contrast to the inferential mindreading proposed by theories such as TT and ST. Commonly, DSP theorists argue that we have direct access to the basic emotional states and motor intentions of other people (Gallagher, 2008; Gallagher, 2015; Krueger and Overgaard, 2012; Newen, Welpinghus, and Juckel, 2015; Spaulding, 2015; Zahavi, 2011). On these accounts, perceiving the emotional state of another person is on a par

5 This can be seen as a response to Bohl (2015) who raises the worry for DSP theorists that they have not specified in which contexts DSP rather than other mindreading strategies are used.

6 I wish to clarify two readings of this statement that I do not intend. The first reading is that DSP is *only* used in close relations, and hence not used to recognize the emotions of strangers. This is false, since we can directly perceive basic emotions of strangers as well. The second reading is that in close relations, we only use DSP, as opposed to any other mindreading strategies. Although I find this more plausible, it is likely that we sometimes have to resort to other strategies as well, for example when we are not able to directly perceive that a partner is hurt without them saying so. My main claim in this paper is that we most often use DSP, as opposed to other mindreading strategies, in close relations. Thanks to an anonymous referee for pushing me to clarify this.

7 See Spaulding (2015) for an account of DSP whereby only occurrent basic emotions and not dispositional emotions are perceived.

with perceiving properties such as being a table or a chair (Siegel, 2006).[8] When perceiving the property of being a table, there is no extra step of inferring or judging that it is a table. Similarly, when I see a person smiling, I do not need to infer that they are happy. Rather, I see the happiness directly in their expression since the emotion is embodied in the expression.
It is important to note that DSP is not the same as behaviorism.[9] Whereas behaviorism claims that emotional states can be *reduced* to expressive states, DSP claims that *expressions partly constitute emotional states*.[10] Thus, happiness cannot be reduced simply to smiling, but rather smiling is a constitutive part of happiness.[11] Other parts include physiological responses (such as increased heart rate), phenomenal experience, cognitive features (such as attitudes and shifts of attention and perception), and an intentional object (Newen, Welpinghus, and Juckel, 2015: pp. 192-193).[12]
DSP does not need to be committed to the view that emotional expressions are necessarily partly constitutive of emotional states – one could be sad without crying. When this is the case, a pluralist theorist could claim that we cannot *directly* access the emotional state of the other person since it is not expressed, but it is still possible to access it via an inferentialist strategy. However, there are good reasons to think that these emotions in many cases are directly perceived, as there are characteristic expressions of embodied emotions that we readily recognize. There has been evidence showing that we can identify both emotional states and intentions from perceiving someone's posture or the way in which an action is executed.[13] Studies have shown that we can correctly categorize emotions when seeing point-light or full-light displays of both moving and static bodies (Atkinson *et al*., 2004; Heberlein *et al*., 2004). In a full-light display, an actor's body – but not facial expression – is visible to a subject, whereas in a point-light display, the actor wears luminescent straps on selected parts of their body (such as wrists, shoulders and knees) such that only the straps are visible as light points. In both these cases, subjects were able to tell what emotion the actor was expressing.
Further, there is a close relationship between the expression and the phenomenal experience of emotions, such that our emotional experiences are also affected by our expressions. This lends support to the idea of embodied emotions. For example, studies of people with Moebius syndrome support that facial expressions affect the phenomenology of emotions. Moebius syndrome is a congenital condition whereby a person's facial muscles are paralyzed and their eyes cannot move laterally. According to verbal accounts from people with Moebius syndrome, some feel like they "liv[e] in their head" and do not recognize themselves as *feeling* happy or *feeling* sad, but rather just as "think[ing] happy or think[ing] sad" (Cole, 1998). It seems as though they do not experience the phenomenology of happiness, and part of the reason for this could be that they are unable to express happiness by smiling. Turning to how emotional states are expressed by bodily actions, studies have also found that the intensity of the emotional experience is reduced when people are not able to express their emotion using body language. This can be seen in individuals who have suffered spinal cord injuries (Chwalisz *et al*., 1988; Hohmann, 1966; Laird, 2007; Mack *et al*., 2005). Studies have also

8 Spaulding characterizes the directness with which we perceive emotions to be on a par with how we perceive objects, rather than properties. See Spaulding (2015, p. 3).
9 See Krueger and Overgaard (2012) for an argument against DSP being a kind of behaviorism.
10 Whether or not DSP needs expressions to be partly constitutive of emotions might be a controversial issue, though it is usually assumed and argued for in the literature (Krueger and Overgaard, 2012). Here, I will set this issue aside since it is outside the scope of this paper. Thanks to an anonymous reviewer for bringing this to my attention.
11 I am here using 'constitutive part' in the same sense as Green (2010) and Krueger and Overgaard (2012).
12 I will not argue for the inclusion of these features here. For discussion, see Newen, Welpinghus, and Juckel (2015).
13 For work on decoding intentions from movement kinematics, see Ansuini *et al*. (2016); Cavallo *et al*. (2016).

manipulated facial expressions, posture, and gestures to produce a corresponding change in emotional phenomenology (Davis *et al*., 2009; Laird, 2007; Niedenthal, 2007). These studies suggest that there is a causal link between the emotional expression, the physiology and the phenomenology of the emotion. Such a tight causal link is a good reason to think of these features as making up a distinct system and, as such, being constitutive parts of that system. We would readily take this to be the case when considering other systems, such as a computer. In this case, there are causal connections between the internal working parts, such that the activation of one will cause the activation of another, and we further take all the internal parts to be constitutive of the computer as a whole. Causality and partial constitution are thus not mutually exclusive here.

Let us now turn to the question of how we can perceive the emotion as an entity by perceiving one part of it. There is a worry that we cannot perceive others' emotions since we cannot perceive the physiological or cognitive states of another's emotion. A response to this remark is that we can perceive entities by perceiving characteristic parts of them (Green, 2010; Krueger and Overgaard, 2012). Expressions are characteristic parts of emotions, and perceiving a characteristic part of an entity is sufficient for perceiving that entity.[14] Bodily expression is part of an emotion, e.g. crying is part of being sad. By perceiving someone crying, we can thus perceive their emotional state of sadness.

There are two worries pulling in opposite directions that one might have with regards to perceiving emotional states by perceiving characteristic expressions. The first worry is that one might think that since emotions are embodied, we should be able to perceive any and all emotions in anyone. This does not seem to be the case, as only basic emotions are perceived cross-culturally (Ekman, Friesen, and Ellsworth, 1972). The second worry is that we should only be able to perceive certain emotional states, since not all emotional states have characteristic expressions. Both of these worries can be resolved by clarifying what 'characteristic' means. It is useful to draw a distinction between perceiving the emotions of strangers by perceiving *characteristic expressions of basic emotions*, and perceiving the non-basic emotions of people we are close to by perceiving those *particular people's characteristic expressions*. In the case of strangers, since we have little exposure to their expressions, we are only able to directly perceive basic emotions that are individuated in characteristic ways, e.g. smiling to express happiness. Empirical research has shown that we are able to recognize these expressions both within our culture and cross-culturally, lending support to the idea that these expressions are universal (Ekman, Friesen, and Ellsworth, 1972). 'Characteristic' here thus refers to what is generally (or universally) characteristic of an emotion for *any* given person. With a person we are close to, however, the expression need not be characteristic of how people *in general* express the emotion, but only of how *that person in particular* expresses the emotion (this has been dubbed *individual-typical expressions* – Glazer, forthcoming). If I know a person who always expresses surprise by frowning, I will come to recognize that, *for that person in particular*, frowning is a characteristic component of surprise. How we learn this will be clarified in Section 4. In the next section, I will argue that we also need a view of perceptual content as conceptually rich in order for mindreading to be a perceptual process.

## 2. Rich Perception

This section will focus on Siegel's account of rich perceptual content, and how we come to acquire the concepts necessary to perceive e.g. anger *as anger*, rather than just as a pattern of recognizable features. Rich perception is the idea that the content of perception is not limited

14 I will not dwell on this point since it has been extensively discussed by Green (2010) and Krueger and Overgaard (2012). See these papers for more detailed arguments.

to representing low-level features such as shapes and colors, but is instead extended to high-level features which can include emotions and intentions. Siegel argues that, in perception, we also represent properties such as (but not limited to) 'being a natural kind'.[15] Thus, we perceive a tiger *as a tiger*, rather than as a bundle of features which are inferred be a tiger.

Siegel's argument runs from the phenomenology of perception to the representational content of the visual phenomenal experience. Suppose that you have never seen a pine tree before, and are hired to cut down all the pine trees in a forest containing trees of various kinds. It is pointed out to you what pine trees look like. After some time, your ability to tell pine trees from other trees improves, and pine trees gradually become more salient to you. Siegel claims that there is a phenomenological difference in the visual experience before and after you were able to successfully pick out pine trees. This phenomenological difference is due to a sensory difference between the two experiences, which is in turn due to the fact that the two experiences differ in content. That is, the content of the visual experience before learning what trees are pine trees is different to the content of the visual experience after having learnt which trees are pine trees.[16] The best explanation for this change in experience is that pine trees are represented *as the kind 'pine tree'* when we become familiar with them. Since all the low-level properties, such as the color and the shape of the trees, are the same in both experiences, low-level representational content is not enough to explain the phenomenal contrast before and after acquiring the concept PINE TREE. That is, improved pattern recognition of low-level features is not enough to account for the different ways in which we experience a pine tree before we are familiar with it and after we are familiar with it; the concept PINE TREE needs to be acquired for the phenomenological shift to occur. I will now turn to the issue of how we come to recognize the pattern of features that make up a pine tree, since this is important for acquiring the concept.

In my view, we come to acquire the concept partly *by means of* becoming better at picking out certain features, i.e. a pattern. In order to see how emotional concepts get into the content of perception, I draw on Wu's argument concerning the relationship between action and attention, and how to solve the Many-Many Problem (Wu, 2008; 2011). The Many-Many Problem is the problem of how we decide to perform a particular action (e.g. cutting down a pine tree), given that we are first faced with many perceptual inputs and need to decide which are relevant to our goal, and then faced with many possible actions that can be performed to reach the goal. By using intention to guide attention, we can solve this problem and focus on the relevant perceptual features of the world in order to act in the best way possible. An agent must first locate the attention to select a specific target, e.g. a glass that they wish to drink from. They must then parse the attention to focus on the specific properties of the object that are relevant to their goal. It is not relevant whether the glass has a particular color, but it is relevant whether it has a handle since that renders it more pick-up-able. Wu argues that concepts play a role in the attentional parsing of a glass. That is, lacking the concept GLASS, I would not be able to attend to the features relevant for my using the glass.[17] Acquiring the concept GLASS allows me to see the glass *as a glass*. There is a clear link here to Siegel's example of seeing a pine tree *as a pine tree* and its features then becoming more salient. When it is pointed out to me what a pine tree looks like, and I have the intention of cutting down all pine trees, my intention can guide my attention such that certain characteristic features of

15 See Siegel (2006; 2010) for a full defense of this argument.

16 See the literature on perceptual learning for empirical evidence supporting this thesis (e.g. Connolly, 2017).

17 Wu allows for the possibility that we can still accidentally use the glass in the right way (Wu, 2008, p. 1019).

the pine tree become more salient to me.[18] My skill of recognizing pine trees thus improves as I become better at picking out the pine tree pattern, allowing me to acquire the concept PINE TREE.

In what follows, I will give an account of the pattern recognition we, as humans, develop. I will consider why we should take this to be a conceptual process rather than a non-conceptual one, by looking at an account that holds pattern recognition to be non-conceptual.

## 3. Against Non-Conceptual Recognition of Emotions

Newen and Schlicht developed a non-conceptual account of pattern recognition of emotions, and Newen, Welpinghus, and Juckel argued that we recognize emotions by recognizing individuated patterns of emotional states (Newen and Schlicht, 2009; Newen, Welpinghus, and Juckel, 2015). I agree with Newen, Welpinghus, and Juckel's characterization of pattern recognition of emotions, but I will argue that we *do* require the concept of e.g. anger to recognize the expressive pattern *as anger*.

According to Newen, Welpinghus, and Juckel, emotional states are individuated as patterns of characteristic features. This can most easily be seen in basic emotions which are said to be recognized cross-culturally (Ekman, Friesen, and Ellsworth, 1972). Plausibly, the reason why they are recognized cross-culturally is because they are all individuated as patterns of characteristic features (which may or may not be hardwired). In any given case of pattern recognition, Newen, Welpinghus, and Juckel claim that a feature F (such as a facial expression) is constitutive of a pattern P if it is part of at least one set of features which is minimally sufficient for a token pattern to belong to type T. The minimally sufficient conditions are then jointly sufficient for the emotional episode to be classified as e.g. the type anger. If one or more of these features were removed, the token emotional episode would no longer be classified as one of anger. It is plausible that we are able to develop this reliable way of recognizing emotional patterns because we are very often exposed to these particular emotional expressions. The mechanism thus has a large database of expressions to draw on in order to establish the characteristic patterns of basic emotions. In the final section, I will show how this feature of pattern recognition is also important for recognizing non-basic emotions in people we are close to.

Before doing so, let us discuss why Newen and Schlicht's claim that there can be a non-conceptual understanding of others' emotions is to be rejected (Newen and Schlicht, 2009, p. 232). This understanding is supposed to be underpinned by mirror neuron processes.[19] Mirror neurons are neurons that fire both when an action is executed by an agent, and when the agent sees another executing the same action. For emotions, research has shown that the same brain areas are activated when a person feels disgust, as when they see someone else looking disgusted (Wicker *et al.*, 2003). The thought is that we can non-conceptually perceive another person's emotional state using the mirror neuron system.[20] However, as argued above, perception is a conceptual process and the understanding of emotions cannot be explained only in terms of non-conceptual processes. To further strengthen this point, I will show that studies on subjects with dementia support that concepts are necessary both for using objects correctly and for recognizing emotions.[21] These similarities also support my application of Wu's account of attention to emotion recognition.

18 Importantly, the first intention cannot involve the concept PINE TREE, since this has not been learnt yet. Instead, I suggest that the content of the intention would be something like 'to cut down all trees looking thusly', where 'thusly' refers to whatever features one's instructor pointed out.

19 For work on mirror neurons, see Gallese *et al.*, 1996; Iacoboni, 2005.

20 Mirror neurons have often been taken as providing evidence in favor of Simulation Theory (Goldman and Gallese, 1998; Goldman, 2006). Newen and Schlicht (2009) argue against this.

21 Wu makes a similar argument to show that concepts are necessary for correct use of objects (Wu, 2008).

In order to possess concepts, semantic memory is needed to store these concepts. It is thus plausible to think that a person with impaired semantic memory would lose some of their concepts, and therefore not be able to use the objects corresponding to these concepts. This is indeed the case in semantic dementia, as demonstrated in a study by Hodges (2000). Subjects were tested on their ability to use everyday objects, as well as naming the objects and associating them with their use (e.g. a cork screw is used to open a bottle). To test whether the subjects could use the objects, the subjects' grasp, orientation and movement were tested. Failure to use the object in a correct manner correlated with semantic disabilities. In experiments testing subjects without dementia, it has also been shown that their grasping of objects is impaired when having to deploy their semantic memory in a distractor task (Creem and Proffitt, 2001). The claim that concepts are needed for emotion recognition can be supported in the same way. Studies have shown that impaired semantic memory also has an impact on emotion recognition. Subjects with dementia were tested on emotion recognition, and performed significantly worse than control groups without dementia (Keane *et al*., 2002; Lavenu *et al*., 1999; Rosen *et al*., 2004 ). This supports that the perceiving of emotions is indeed conceptual if we are to see an emotion *as a particular emotion*. That is, in order to see sadness *as sadness*, an agent needs to possess the concept SADNESS. I still grant that there could be a non-conceptual way of registering emotions, but it would be wrong to call this non-conceptual *understanding* as Newen and Schlicht do (Newen and Schlicht, 2009, p. 234). In order to understand the emotion of another person, we need to do more than simply recognizing a pattern which allows us to distinguish *this pattern* from *that pattern.* Being able to distinguish a mere pattern of sadness from a mere pattern of anger does not entail that an agent understands what distinguishes the *emotion sadness* from the *emotion anger*. I thus take it that the understanding of emotions using DSP is a conceptual process where concepts are the content of perception. Next, I will tie this together with how we develop a direct perception of complex emotions in close relations.

## 4. Directly Perceiving Complex Emotions in Close Relations

Finally, I wish to show that we can directly perceive complex emotional states of people we know well. In a way, we become experts at recognizing the individuated patterns of emotions in close relations.[22] Again, I will draw on Wu's account of attention to show how we can come to recognize these more complex patterns.

First, *prima facie* it seems like we can recognize a wider range of emotions in people that we are close to, such as spouses, family or friends.[23] I will focus the discussion around friendship, but my account extends to any close relationship. When seeing a stranger's smile, you might recognize the basic emotion it expresses (happiness), but you will not be able to distinguish that way of smiling from how the person smiles when proud; your pattern recognition is not fine-grained enough. This is because you are not familiar with that person's *particular expressions*, i.e. how *they in particular* smile when being happy as opposed to proud. You are able, however, to recognize particular ways in which a close friend smiles. You can easily distinguish how they smile from being bemused from how they smile from being happy or proud. That is, your skill at recognizing their particular expressions has improved. I will show how intention-guided attention plays a crucial role in this.

It could be thought that our pattern recognition improves only in virtue of *exposure* to the friend. This seems like a plausible suggestion since we tend to spend more time with friends

22 That non-conceptual mindreading is used in family relations is suggested by Newen and Schlicht (2009) in a footnote, but there is no further development of this account.

23 To my knowledge, there has not yet been any research on emotion recognition in close relationships, but there is research showing that we are better at recognizing emotions in cultures we are more frequently exposed to. This fits in well with my account. See Elfenbein and Ambady (2003) and Elfenbein *et al*. (2007).

than with other people. However, consider the following case. Samira works as a vet. In commuting to work, Samira is every day exposed to the same people on the tram. She is thus every day exposed to these people's expressions, and one might think that this is all that is needed for her to be able to develop direct perception of their complex emotions. Still, I think there is a good reason to think that her pattern recognition and ability to directly perceive the emotions of her fellow commuters are restricted in this case, and do not improve further. This is, at least in part, because Samira lacks the right intention in these cases.[24] As with the pine trees, an intention is needed to guide the attention to pick out the patterns relevant to a complex emotion. Normally, we do not have the intention to get to know our fellow commuters, and therefore our attention is not guided towards improving pattern recognition. However, if Samira were particularly interested in a person she saw on the commute every day, and had the intention of getting to know them, it is possible that she would come to be able to directly perceive some emotional states of that person. It is unlikely that she would come to be able to directly perceive many complex emotional states, however, since she is only exposed to the person in one context. If she were exposed to the person in more contexts – such as during a family dinner, when receiving some good news, when a friend makes an unfair remark, etc. – she would likely experience a wider range of that person's emotions. The range of emotions expressed by a person during a commute is likely to be smaller since the context is somehow monotonous.

In friendship, on the other hand, we are exposed to the friend in many contexts, and we do have the intention to e.g. get to know them better. It is important to note that the intention here is not as explicit as 'to learn to read the facial expressions better', but neither is the intention in Wu's example of picking up a cup 'to pick up the cup in the best way possible' – rather, the intention is just 'to be a good friend' or 'to drink from the cup'. Since part of what it means to be a good friend is to be able to respond to the other person's emotional states in an appropriate way, intending to be a good friend will guide attention in the right way for direct social perception to develop. Since we are exposed to a wide range of emotions in close relations, we can also come to directly perceive complex emotions such as jealousy. This follows directly from my claims about pattern recognition. Basic emotions are individuated by patterns that can be recognized cross-culturally, whereas jealousy might be expressed differently in different people, but crucially in *characteristic ways in particular people*. Having observed jealousy in a friend on multiple occasions, I can learn to recognize that pattern in the same way as I can recognize happiness in most strangers on the street. We can thus directly perceive the basic emotions of most people, but we additionally directly perceive complex emotions of people we are close to.

Finally, once we are able to perceive complex emotions of people we are close to, other mindreading strategies become redundant in face-to-face interaction. If I can directly perceive that my friend is proud, it is redundant to use another mindreading strategy that requires more steps. Since emotions are embodied, the information needed is already there in perception. Other mindreading strategies are still needed in close relations in other cases, such as if a friend texts me. Since I am not able to directly perceive their emotional state from the words I am reading, I will have to resort to another mindreading strategy.[25]

24 The case is probably more complicated than as outlined here. The way in which we interact with people we are close to – e.g. how we confide in each other and how we openly avow feelings that we might otherwise keep to ourselves – could also contribute to the development of direct perception. My thanks to an anonymous reviewer for pointing this out.

25 This account could be seen as an explanation for why we also *feel* closer to people that we know well. It seems plausible that actually having direct perception of someone's mental states might also affect your relationship with

## 5. Conclusion

In this paper, I provided an account of DSP, which hinges on the ideas of rich perception, attention, and pattern recognition. I first distinguished DSP from other accounts of mindreading, such as TT and ST, claiming that the main difference between these camps is that DSP is a non-inferential kind of mindreading. I then argued that emotional expressions partly constitute emotional states. For DSP to be a viable theory, we need to be able to perceive emotional states by seeing emotional expressions. I supported this by showing that emotional expressions affect the phenomenology of emotions. For example, people with Moebius syndrome or people who are paralyzed and therefore not able to bodily express their emotions, experience emotions differently. Expressions should thus be considered as partly constitutive of emotional states. Since we can perceive an entity by perceiving a part of it, I suggested that, by perceiving part of an emotion (i.e. its expression), we can perceive the emotion. I then considered the nature of perception, and how emotions can be part of perceptual states. Siegel argues that concepts feature in our perception, and that there is a contrast between experiencing a pine tree before having acquired the concept PINE TREE and experiencing a pine tree after having acquired the concept. Certain features of the pine tree become salient to us when we learn the concept, and we become able to recognize the pine tree *as a pine tree.* Similarly, when studying someone's face, we learn to recognize the patterns that individuate particular emotions. It is necessary both that we learn this pattern and that we possess the concept of a particular emotion, such as anger, to recognize it *as anger*. Drawing on studies which show that people with dementia are not able to recognize emotions for which they lack a concept, I maintained that the perception of emotions is a conceptual process, contra Newen and Schlicht. Finally, I argued that DSP is particularly important in close relations. Adopting Wu's theory of attention, I showed how we can come to recognize complex emotions of people that we are close to by intention-guided attention. In virtue of intending to get to know a person better, certain features become salient to us such that we are able to recognize these as patterns of complex emotions. Theoretical considerations suggest that DSP is particularly important for close relationships, but future empirical research is also needed to validate this claim.

### REFERENCES

Ansuini, C., Cavallo, A., Campus, C., Quarona, D., Koul, A., & Becchio, C. (2016). Are we real when we fake? Attunement to object weight in natural and pantomimed grasping movements. *Frontiers in Human Neuroscience*, 10(471).

Atkinson, A., Dittrich, W., Gammell, A, and Young, A. (2004). Emotion perception from dynamic and static body expressions in point-light and full-light displays. *Perception*, 33, 717-746.

Bohl, V. (2015). Continuing debates on direct social perception: Some notes on Gallagher's analysis of "the new hybrids". *Consciousness and Cognition*, 36, 466-471.

Borg, E. (2007). If mirror neurons are the answer, what was the question?. *Journal of Consciousness Studies*, 14, 5-19;

Cavallo, A., Koul, A., Ansuini, C., Capozzi, F., & Becchio, C. (2016). Decoding intentions from movement kinematics. *Scientific reports*, 6(37036), 1-8.

Chwalisz, K., Diener, E., & Gallagher, D. (1988). Automatic arousal feedback and emotional experience: evidence from spinal cord injured. *Journal of Personality and Social Psychology*, 54(5), 820-828.

Cole, J. (2010). Agency with impairments of movement. In D. Schmicking and S. Gallagher (Eds.), *Handbook of Phenomenology and Cognitive Science.* Dordrecht: Springer, 655-670.

them such that you feel like you have a closer connection to them. This would need to be further developed in another paper.

Colombetti, G., & Roberts, T. (2015). Extending the extended mind: the case for extended affectivity. *Philosophical Studies*, 172(5), 1243-1263.
Connolly, K. (2017). Perceptual Learning. *The Stanford Encyclopedia of Philosophy*, E. N. Zalta (Ed.). https://plato.stanford.edu/archives/sum2017/entries/perceptual-learning.
Creem, S., Proffitt, D. (2001). Grasping objects by their handles: A necessary interaction between cognition and action. *Journal of Experimental Philosophy: Human Perception and Performance,* 27, 218-228.
Davis, J., Senghas, A., & Ochsner, K. (2009). How does facial feedback modulate emotional experience?. *Journal of Research in Personality*, 43(5), 822-829.
Ekman, P., Friesen, W.V., & Ellsworth, P. (1972). *Emotions in the Human Face*. Oxford: Pergamon Press.
Elfenbein, H.A., & Ambady, N. (2003). When familiarity breeds accuracy: Cultural exposure and facial emotion recognition. *Journal of Personality and Social Psychology*, 85, 276-290.
Elfenbein, H.A., Beaupré, M., Lévesque, M., & Hess, U. (2007). Toward a dialect theory: Cultural differences in the expression and recognition of posed facial expressions. *Emotion*, 7(1), 131-146.
Fiebich, A., & Coltheart, M. (2015). Various ways to understand other minds: Towards a pluralistic approach to the explanation of social understanding. *Mind and Language*, 30(3), 235-258.
Fiebich, A., Gallagher, S., & Hutto, D. (2016). Pluralism, interaction and the ontogeny of social cognition. In J. Kiverstein (Ed.), *The Routledge Handbook Philosophy of the Social Mind*. London: Routledge, 208-221.
Gallagher, S. (2008). Direct perception in the intersubjective context. *Consciousness and Cognition*, 17(2), 535-543.
Gallagher, S. (2015). The new hybrids: Continuing debates on social perception. *Consciousness and Cognition,* 36, 452-465.
Gallagher, S., & Hutto, D. (2008). Understanding others through primary interaction and narrative practice. In J. Zlatev, T. Racine, C. Sinha & E. Itkonen (Eds.), *The Shared Mind: Perspectives on Intersubjectivity*. Amsterdam: John Benjamins Publishing Company, 17-38.
Gallese, V., Fadiga, L., Fogassi, L., & Rizzolatti, G. (1996). Action recognition in the premotor cortex. *Brain*, 119, 593-609.
Gallese, V., & Goldman, A. (1998). Mirror neurons and the simulation theory of mind-reading. *Trends in Cognitive Sciences*, 2(12), 493-501.
Glazer, T. (forthcoming). Looking angry and sounding sad: The perceptual analysis of emotional expression. *Synthese*, 1-25.
Goldman, A. (2006). *Simulating Minds: The Philosophy, Psychology, and Neuroscience of Mindreading.* New York: Oxford University Press.
Gopnik, A., & Wellman, H.M. (1992). Why the child's theory of mind really is a theory. *Mind and Language*, 7 (1-2), 145-171.
Green, M. (2010). Perceiving Emotions. *Aristotelian Society Supplementary,* 84(1), 45-61.
Heberlein, A., Adolphs, R., Tranel, D., & Damasio, H. (2004). Cortical regions for judgements of emotions and personality traits from point-light walkers. *Journal of Cognitive Neuroscience*, 16(7), 1143-1158.
Hodges, J., Bozeat, S., Lambon Ralph, M., Patterson, K., Spatt, J. (2000). The role of conceptual knowledge in object use: Evidence from Semantic Dementia. *Brain,* 123, 1913-1925.
Hohmann, G.W. (1966). Some effects of spinal cord lesion on experienced emotional feelings. *Psychophysiology*, 3(2), 143-156.
Iacoboni, M., Molnar-Szakacs, I., Gallese, V., Buccino, G., Mazziotta, J.C., & Rizzolatti, G. (2005). Grasping the intentions of others with one's own mirror neuron system. *PLOS Biology*, 3(3), 529-535.

Keane, J., Calder, A., J., Hodges, J. R., & Young, A.W. (2002). Face and emotion processing in frontal variant frontotemporal dementia. *Neuropsychologia*, 40(6), 655-665.
Krueger, J., Overgaard, S. (2012). Seeing subjectivity: defending a perceptual account of other minds. *ProtoSociology*, 47, 239-262.
Laird, J.D. (2007). *Feelings: The perception of self.* Oxford: Oxford University Press.
Lavenu, I., Pasquier, F. (2005). Perception of emotion on faces in frontotemporal Dementia and Alzheimer's Disease: A longitudinal study. *Dementia and Geriatric Cognitive Disorders*, 19, 37-41.
Mack, H., Birbaumer, N., Kaps, H., Badke, A., & Kaiser, J. (2005). Motion and emotion: Emotion processing in quadriplegic patients and athletes. *Zeitschrift Fur Medizinische Psychologie*, 14(4), 159-166.
Newen, A., & Schlicht, T. (2009). Understanding other minds: A criticism of Goldman's simulation theory and an outline of the person model theory. *Grazer Philosophische Studien*, 79(1), 209-242.
Newen, A., Welpinghus, A., & Juckel, G. (2015). Emotion recognition as pattern recognition: The relevance of perception. *Mind and Language*, 30(2), 187-208.
Niedenthal, P. (2007). Embodying emotion. *Science*, 316(5827), 1002-1005.
Rosen, H., Pace-Savitsky, K., Perry, R., Kramer, J., Miller, B., & Levenson, R. (2004). Recognition of emotion in the frontal and temporal variants of frontotemporal Dementia. *Dementia and Geriatric Cognitive Disorders*, 17, 277-281.
Siegel, S. (2006). Which properties are represented in perception? In T. S. Gendler and J. Hawthorne (Eds.), *Perceptual Experience*. New York: Oxford University Press, 418-503.
Siegel, S. (2010). *The Contents of Visual Experience*. New York: Oxford University Press.
Spaulding, S. (2015). On Direct Social Perception. *Consciousness and Cognition*, 36, 472-482.
Wicker, B., Keysers, K., Plailly, J., Royet, P., Gallese, V., & Rizzolatti, G. (2003). Both of us disgusted in my insula: The common neural basis of seeing and feeling disgust. *Neuron*, 40(3), 655-666.
Wieser, M., & Brosch, T. (2012). Faces in context: a review and systematization of contextual influences on affective face processing. *Frontiers in Psychology*, 3(471).
Wu, W. (2008). Visual attention, conceptual content, and doing it right. *Mind*, 117(468), 1003-1033.
Wu, W. (2011). Confronting the Many-Many Problem: Attention and agentive control. *Noûs*, 45(1), 50-76.
Zahavi, D. (2011). Empathy and Direct Social Perception: A phenomenological proposal. *Review of Philosophy and Psychology*, 2(3), 541-558.