Radboud Universiteit Nijmegen

Faculty of Philosophy, Theology and Religious sciences

Phenomenological simulation

Masterthesis

Author:
Thijs klein Baltink, s4359763

Supervisor:
Nina de Boer

*Words:*15650 *Date:*28-05-2023

Thesis for obtaining a "Master of arts" degree in philosophy Radboud University Nijmegen

I hereby declare and assure that I, Thijs Bastiaan klein Baltink, have drafted this thesis independently, that no other sources and/or means other than those mentioned have been used and that the passages of which the text content or meaning originates in other works - including electronic media - have been identified and the sources clearly stated. Place: Nijmegen date: 28-05-2023

Contents

Abstract

This thesis seeks to determine if a combination of simulation theory and direct perception can be achieved which unifies them while preserving their unique explanatory advantages. I analyze differences and criticisms between these approaches, and how these apply to a combined theory. I argue how a combined theory which preserves the unique perspectives sidesteps these criticisms. Simultaneously, how simulation and direct perception might inform each other is explored, allowing for complementary action between them. This results in an approach balancing between an integrative hybrid theory, which risks overlooking unique explanatory advantages, and a segregating pluralist view, which fails to address interactions between the theories that could enhance our understanding of social cognition.

1 Introduction

Furthering our understanding of cognition, of our experiences, senses, and how we think, has been a primary goal of neuroscience and philosophy of mind. A subsection thereof aims to understand *social* cognition[1]; explaining how we understand others, our interactions with them, and understanding others' intentions. The capacity for such understanding is historically called 'mindreading' or 'mentalising'; the ability to understand the mental states of others. Social interactions and understanding thereof play a key role in our everyday life. Would we still be able to have meaningful interactions with others or form meaningful connections with them? How could we function as a society if we, as individuals, did not have some insight into the minds and emotions of those around us? How could we work together *without* being able to understand, predict, and rely on the thoughts and emotions of the people close to us? Such understanding is necessary for grand cooperation like the building of the pyramids, the development and deployment of life-saving polio inoculations, and maintaining society.

How has philosophy of mind tried to explain social cognition? This has been attempted in many ways. The common-sense understanding of other minds is often called 'folk psychology'. Early on the idea arose that we need a 'theory of mind', some theoretical understanding of the minds of others. The idea of viewing folk psychological understanding as such a theory of mind is called *theory theory*(Lewis 1970; Dennett 1985). Theory theory claims that we understand other people's mental states, and are able to predict their behaviour, by referencing what we observe and know of them to some rule-based theoretical framework we develop of other minds. Let's briefly illustrate how theory theory would understand the sadness of a child and is able to predict that she will cry. Imagine you see a smiling child walking by with an ice cream cone, suddenly she stumbles and drops the entire thing. Even before she cries you understand that she is saddened by this. In this example, theory theory would argue that you understand the girl's mental state by referencing what you know of her situation (she accidentally dropped her ice cream), against some theoretical rule (from

[1] The definition of social cognition is rather ambiguous, with sometimes authors specifying definitions of their own, and many authors leaving the interpretation open-ended. For this thesis 'social cognition' can be interpreted in a broad way where any cognition about and/or involving others is concerned.

your theory of mind). In this case, it could be the folk psychological concept imparted upon you early in life that children like ice cream and would be saddened by its sudden loss, often crying when experiencing such emotion.

Alternatively *simulation theory*(Gordon, 1986) arose as an explanation for how we understand others. Growing in popularity after empirical discoveries on mirror neurons and the role they might play in social cognition, which was argued to align with a simulation theorist explanation(Gallese & Goldman 1998; Shanton & Goldman 2010). Simulation theory stipulates that we can attribute mental states to others and predict them through simulation; that is to say, we take what we observe and know of the other and feed this into an internal simulation through which we predict what the other feels or how they will act. Understanding the child from a simulation theorist perspective would have you simulate what it would be like for the child to experience this, and predict and understand her through this simulation. You would substitute what little you know of the child's context for your own. Imagine you are a young child, happily walking about with your ice cream. You stumble and drop it. How would *you* feel? You would feel sadness, and you attribute this feeling to the child whose situation you have simulated.

Later, phenomenologically inspired explanations of social cognition arose such as through 4E cognition. 4E approaches to social cognition describe a broad category of theories which look at the embodied; embedded; enacted; and extended properties of social cognition. The phenomenological tradition it draws inspiration from aims to further our understanding of the world based on how we experience it. 4E reasons phenomenologically in that they argue for these various properties of social cognition on the basis of how we experience it; not as some sort of introspective referential process. The primary theory this thesis discuss can be said to be both phenomenological and an early example of 4E; Gallagher's work on *direct perception*(Gallagher 2001; 2008; 2015; 2020). Gallagher argues that (most of) our social understanding comes from being able to directly perceive other's mental states. Self-described as embodied, enacted, and offering phenomenological arguments to ground this account of social cognition, direct perception views social cognition less so as 'mindreading' than as 'bodyreading'(Gallagher 2001; 2007). Direct perception would argue that we understand the child's sadness because we can plainly perceive it in the child and her actions. She embodies the sadness in her expressions and motions, understandable for us without a need to rely on some internal process of inference.

So, on the one hand, we have introspective approaches to social cognition in theory theory and simulation theory, which both look to internal reasoning for explaining our understanding of others. On the other hand, we have phenomenological approaches in direct perception and broader 4E, who look outwardly instead. Might these approaches be united somehow, and could such a combined theory result in a more accurate explanation of social cognition? Direct perception offers great insight into our most common experiences of social cognition, whereas simulation theory can provide an understanding of those parts of social cognition which do not seem directly observable, might a combined theory provide both these advantages?

Beyond the potential for a theory which helps us better understand social cognition, what makes a closer look at combining these different perspectives relevant? Social cogni-

tion permeates our everyday lives and being able to understand others is essential to them. Having a greater understanding of our social cognition means a greater understanding of the human experience. Beyond the general epistemological relevance of furthering knowledge on a subject so intimately tied into our experiences; however, there is good reason for further exploration of a combination of simulation theory and direct perception in particular.

Both classically introspective takes on social cognition (such as simulation theory) as well as phenomenologically reasoned approaches (such as direct perception) argue that they better explain certain aspects of social cognition. A combination has the potential to achieve both explanatory advantages and thus should be explored. But why look at simulation theory and direct perception, when there are plenty more theories stemming from these perspectives? The answer to this question is four-sided. First off, Gallagher's direct perception reasons phenomenologically and is self-described as both embodied and enacted(Gallagher 2001; 2007; 2008) making it a solid choice for a phenomenology-inspired take on social cognition and an early example of the 4E realm. On the other hand, we have simulation theory, a popular account of social cognition from the introspective side of the discussion. Secondly, there is a lot of criticism between direct perception and simulation theory, particularly Gallagher has written against simulation theory in favour of his direct perception (Gallagher 2001; 2007; 2008). This provides a focal point for comparing these approaches and plenty of criticism against which to test a combination of these theories. Thirdly, the criticism that simulation theory held against theory theory, and which it used to defend itself against counterarguments stemming from theory theorists, appears very similar to criticism that Gallagher held against introspective theories as a whole (Gordon 1992, Goldman 1992, Gallagher 2001). Furthermore, there are indicators that simulation theory could be considered a sensory-motor process, much like direct perception, or that simulation might be compatible with certain views on embodied cognition (Gallagher 2001, Goldman 2012). This gives reason to believe that the purported differences between simulation theory and direct perception might not be so big as to be insurmountable. Finally, comparing the whole of introspective accounts of social cognition to the phenomenological ones is too broad a scope, and a more precise focus is needed in order to reach any sort of conclusion and take steps advancing this broader topic of interest. Accordingly, this thesis focuses on whether a combination of simulation theory and direct perception, which can maintain the advantages offered by both unique perspectives, is possible.

'Could a combination of simulation theory and direct perception lead to a theory which better explains social cognition than its parts?'

By better explaining social cognition I mean a theory that more accurately explains how our social cognition seems to take place; best aligning with both our empirical evidence and understanding of underlying mechanisms as well as our experiences. Not some theory which explains what would be the most effective means of understanding others, but a theory which better explains what actually occurs in human social cognition. In testing against counterarguments to simulation theory and direct perception I will be able to discover to what extent a combined theory might maintain the advantages of the individual theories, whilst avoiding the faults or disadvantages of both. If it turns out that a significant

portion of the conflict between these theories does not (fully) apply to a combined theory which respects the unique perspectives of its parts, then some combination of the theories must be possible that maintains both sides' claims to fame, providing a chance at a theory through which we can better understand social cognition than its parts.

A potentially controversial topic of discussion among explanations of social cognition relates to developmental claims; when, why, and how certain capacities for social understanding are developed. Much of this discussion originated with 'false belief' tests - tests which try to determine whether someone is capable of attributing beliefs to others that they know to be incorrect, and from what age. Claims surrounding the results of these tests have led to much discussion on how to interpret these results. "Unfortunately this developmental research is extremely difficult to interpret, and I am inclined to agree with S & N (in their Section 4) that it cannot be used to support one theory over the other with any decisiveness."(Goldman 1992) Another controversial topic which sprouted as a result of the developmental discussion was the (in)applicability of mindreading capacities to non-neuro-typical people, particularly people diagnosed with autism. The discussion surrounding these developmental claims is not central to comparing simulation theory and direct perception, despite authors of both sides having weighed in on it. To avoid unnecessarily threading into controversial waters and adding countless caveats, this thesis will focus on the compatibility of a combined theory of direct perception and simulation theory as it would be applicable to neuro-typical adults. This is not to say that I believe it would not be applicable to either children or non-neuro-typicals, on the contrary. However, discussing all the specifics and edge cases that this discussion brings along, as well as having what I view as problematic literature involved, would not be to the benefit of the topic of this thesis.

First I will briefly elaborate on Gordon and Goldman's simulation theory in section 2.1, as well as describe Gallagher's direct perception in 2.2, to more clearly describe the relevant theories and elaborate on the most relevant aspects thereof. In section 3, I will build up towards an illustration of a combined theory. I will analyse the general differences between introspective and phenomenological approaches, and see how these might relate to a view which combines them in section 3.1. Then I will go through the various criticisms between direct perception and simulation theory, aiming to counter these arguments where possible, and see how these might apply to a combined theory instead in 3.2. As well as discuss flaws with previous approaches to combining these perspectives, and how this can be avoided. I will then sketch an idea of what such a combined theory might look like through some examples, showing how such theories could view direct perception and simulation theory as separate but very much complementary tools from our social cognition tool-set in section 3.3. Discussing how these differences and criticisms apply to the combined theory will reveal to what extent such a theory maintains the advantages from these two unique perspectives on social cognition, giving us an answer to the question of this thesis. Finally, in section 4, I will briefly discuss how the explored combination might relate to other projects of incorporating specific theories on social cognition into a larger whole, such as hybrids between theory theory and simulation theory, or the incorporation of direct perception in a broader interaction theory.

The result of this thesis provides reasons for thinking that the conceptual differences

between these approaches need not stand in the way of a combined theory which preserves their unique explanatory advantages, as well as a first notion of what such a theory could look like.

2 Overview of Simulation Theory and Direct Perception

To start off, this section will introduce simulation theory and direct perception. Here I will generally describe how these theories explain social cognition and briefly touch on the more relevant aspects of these theories for the later exploration on how to approach combining them. As such, this section aims to provide a more clear background and frame of reference for the comparisons and argumentation in section 3.

2.1 Simulation Theory

Simulation theory proposes that we understand and predict the behaviour of others by simulating their thoughts, feelings, and actions in our own minds. According to simulation theory, we do this by activating our own mental representations of the mental states of others (e.g., beliefs, desires, intentions) and using those mental representations to understand and predict how others feel and will behave in a given situation. In this sense, simulation is similar to the commonplace concept of 'putting yourself in another's shoes'. That being said the degree to which you put 'yourself' into the other's shoes would be depended on what type of simulation you employ, which I will elaborate on later in this section.

How did simulation theory arise? The idea of social understanding through simulation was first published under 'Folk psychology as Simulation'(Gordon 1986)[2]. This idea was quickly picked up by Goldman (Goldman 1989) and further developed and defended by both authors separately (Gordon 1992; Goldman 1992). However, the descriptions of simulation theory by both authors are very similar, and there is no reason to believe significant incompatibility between their respective accounts. Goldman's initial work on simulation theory is so much in line with Gordon's, that in a work defending simulation theory against critique by theory theorists, Goldman continually lumps his own arguments in with those of Gordon, despite them coming from these separate individual works: "Let me turn now to the seven arguments by Gordon and Goldman (henceforth: G & G)"(Goldman 1992, 106). When looking at the criticism Gallagher has on simulation theory he too refers to Gordon one moment and Goldman the next, without indicating meaningful distinctions between those accounts (Gallagher 2001; 2007; 2008). As such, I too will use 'simulation theory' to refer to the broader work on this theory developed first, and principally, by Gordon and Goldman.

Simulation theory has been very influential in the field of social cognition. It became the predominant theory explaining how we understand others, succeeding the previously prevalent *theory theory*. Some of the initial popularity might be explained through how simulation theory relates well to the experience we have when we consciously try to predict others' intentions. Take Gordon's(Gordon 1986) example of a chess player simulating the point of view of his opponent in order to predict their moves. Much of the later popularity came following empirical evidence of *neural resonance*, such as the emotional mirroring that appears to occur in mirror neurons. These empirical findings were explained by simulation theorists (Gallese & Goldman 1998; Shanton & Goldman 2010), relying on 'lower-level' or

[2]Although there is earlier work opposing theory theory following a similar vein (Morton 1980)

'implicit', i.e. less-than-(fully)-conscious, variants of simulation theory. These simulation theory explanations gained widespread acceptance for the neural resonance phenomenon, as the measured neuronal activity appears to 'mimic' the experiences of others much like the simulation theorist account would[3]. This has cemented simulation theory as a dominant theoretical explanation for social cognition until this day.

We will now look a little more closely at this first work on simulation theory (Gordon 1986), as well as the important elaboration provided by Gordon (1992) which describes the varying degrees of putting yourself in another's shoes I referenced earlier in this section.

Gordon(1986) provides the first account of simulation theory in his article 'Folk Psychology as Simulation'. This terminology might seem odd given the association of folk psychology with theory theory. 'Folk Psychology' at the time was directly associated with the 'theory of mind' people have through which they are capable of understanding others, following the then dominant Theory theory. It is this usage that Gordon refers to, using the title to challenge theory theory's concept of 'Folk psychology as a theory'. In this sense, Gordon(1986) is suggesting simulation as an alternative for explaining social cognition to theory theory. Later works on simulation theory and theory theory prefer the term 'mindreading' to describe the capacity for understanding others. In this article Gordon proposes a novel explanation for understanding how we attribute beliefs to others and how we understand and predict their actions; we do this through simulation, through some sort of 'pretend-play'.

The basic idea behind simulation by Gordon grew first from a specific sort of introspection, or what he describes as "hypothetical-*self*-prediction". This refers to a type of thinking that we often engage in, where we imagine what we would do in a given theoretical scenario. He provides an example of pretending there are sounds of footsteps coming from the basement, and asking yourself the question 'What shall I do now?'. Gordon describes this substitution as a form of pretend play, by pretending certain facts are different than they really are and thinking about how you would act given that situation. He describes the entire exercise ".. a kind of *practical simulation* a simulated deciding *what to do*."(Gordon 1986, 161).

Simulation for the purposes of mindreading; for understanding others, would then be to think similarly, but not with the goal of understanding how we ourselves should act, but understanding how the other will. By putting ourselves in anothers' shoes we can understand their beliefs and mental states, and predict their behaviour. Take Gordon's example of how chess players think similarly by visualising the board from both the opponent's position and their goals for the game. Gordon argues that in *that* simulation the players do not simulate what they themselves are to do in such a theoretical scenario, but what their opponent will likely do in the real world. As such, this early take on simulation is one of 'pretend-play' wherein the observer, mentally, pretends to be the other, simulating what they are thinking and how they will act.

Gordon's projection

A clarification Gordon made on simulation theory lies in what he calls the misconception

[3]Note that this, although a popular explanation, is not without criticism Gallagher's(2008). Such criticism will be discussed in the following section.

that simulation theory *requires* that someone 'puts themselves in another's place'. Instead, Gordon sees simulation as a spectrum, with varying degrees of how much the self is actually put in the other's place. "It is tempting to explain simulation in terms of the familiar notion of 'putting oneself in the other's place'. This would pose a problem for ST, however. We often explain and predict another's behaviour *without* putting ourselves in the other's place– or so it would appear."(Gordon 1992, 12)

To explain this Gordon describes simulation as belonging to either of two types of 'projection'. The 'putting oneself in the other's place', version of simulation would, in this context, be some partial projection, which does not fully project one's own attitudes/environmental awareness/temporal-or-spatial position. Partial projection would instead (attempt to) supplant one or multiple of these aspects (fully or partially) with those of the simulated target. For example, on a hike, you can mentally adjust your own spatial position for that of someone walking ahead of you, who is suddenly alarmed and is telling you to quickly head back. Despite not seeing any apparent danger yourself you could project onto your friend, substituting the spatial dimension, so as to understand that there might well be some terrible danger lurking just around the bend (e.g. a grizzly bear) (Gordon 1992). However, Gordon argues that an adjusted, or partial, projection is not the default approach we take in simulating others, but only occurs whenever some explanation for behaviour turns out to be insufficient. In this case, your friend turning around in alarm is not being understood through a *total* projection, which does not adjust the environmental awareness of your own, and thus cannot imagine the danger ahead. Such *total* projection should not be considered 'putting oneself in the other's place', as it does not adjust the projection for the point of view of its target. This less resource-intensive total projection would be the default *modus operandi* according to Gordon, with adjustments taking place as necessary, resulting in partial projections of varying degrees, to amend a failure in understanding.

Gordon further stresses that this view on simulation theory does not solely (or perhaps even primarily) focus on *predicting* behaviour, but that *explaining* behaviour plays a central role. Wherein a shortcoming of an explanation signals a need to (further) adjust a projection. Such as a high lifting of the legs of a hiker makes little sense from your location, until adjusting for him needing to step onto a large rock makes the explanation fit again (Gordon 1992). But, Gordon repeatedly stresses, the resulting explanation *itself* is also of great value to us. " You see your child climb a tree and pick some fruit, and to account for that behaviour you have to pretend that the fruit isn't (as you know it *is*) poisonous. It is clear what lesson you will have to impart in order to correct this dangerous behaviour."(Gordon 1992, 24) Understanding *why* the child attempts to pick the fruit (lacking awareness that it is poisonous) is what the parent will take away, rather than merely the prediction that the child will attempt to pick the fruit.

Gordon argues it *is* because total projection is unlike simple causal explanations of natural phenomena, which search for theoretical reasons to explain them. Total projection instead looks to *practically* and *emotionally* relevant features to explain someone's behaviour. "When I see a car stop at an intersection, I have a ready explanation if the light is red. Here indeed I am projecting to the driver. But there are also naturalistic explanations in terms of colour: Because the light is red, it doesn't fog the photographic paper in the developing tray. Here I am *not* projecting to the paper." (Gordon 1992, 15). Whether this perspective

stretches the definition of 'simulation' too far will come under further scrutiny in section 3, where I delve into the conflicts between direct perception and simulation theory.

Does simulation have to occur consciously? As we have seen simulation theory provides a framework for understanding how we use our own mind as a model for others' mental states to understand and predict their behaviour. There are various ways in which simulation can occur. It can happen at the 'higher level', sometimes referred to as explicit simulation, which is simulation applied consciously. But it can also happen as a 'lower-level' process, referred to as implicit simulation or a sub-conscious process. Simulation which happens in some grey area between the explicit and implicit, between the conscious and subconscious, can also occur. An analogy is made to such in-between simulation being like driving a car. Various actions are required to successfully drive. A novice driver first consciously shifting gears, whilst after experience having it become a sub-conscious process, yet the process is not fundamentally changed in nature. An important distinction that such a higher-level subconscious process makes, is that when something unexpected happens, for example, your wheels aquaplane while driving, we can still intervene consciously in the same process. Suddenly you might consciously shift gears, consciously steer, and take other actions to avoid an accident. Similarly, simulation theorists argue that when the sub-conscious fails we can take a conscious intervention in the process and adjust the parameters of simulation in an attempt to 'fix' the predictions given. Such intervention would be in line with Gordon's concept of total projection as a default mode of simulation being switched to various degrees of partial projection when the total projection fails to provide understanding.

We are now armed with more background information on simulation theory for our frame of reference. Of particular note is the varying degree of projection which occur in simulation; the amount of 'yourself' which is used. As well as simulation happening as conscious or subconscious processes (or something in between). Storing this away for reference throughout section 3, I will first provide some more background on direct perception.

2.2 Direct Perception

What does 'direct perception' signify? A sensory concept of direct perception originates from Gibson arguing that visual perception is a process wherein people have an innate capacity to directly perceive our environment, rather than experiencing and processing an array of visual sensations or colour-patches (Gibson 1950; 1966). Gallagher's direct perception instead relates specifically to direct perception within the social context, sometimes (more specifically) referred to as 'direct perception in the intersubjective context'(Gallagher 2008) or 'direct social perception'(Gallagher 2015; 2020), yet most often simply as direct perception. "The relevant contrast is not between direct and indirect perception but between perception and something added to perception, e.g., an inference or interpretation that goes beyond what is perceived."(Gallagher 2009, 537). This direct perception is akin to that of Gibson in that it claims we *directly* perceive without inference, but focused on perceiving social qualities, such as someone's emotional or intentional states.

Gallagher argues for direct perception phenomenologically (Gallagher 2001; 2007; 2008; 2020). The phenomenological tradition itself is an older one, dating back to philosophers such as Husserl and Heidegger in the early 20th century and goes far beyond describing a direct perception of socially relevant aspects of the world. However, the recent use of the term within philosophy of mind tends to describe a more specific meaning than this rich tradition, as such it is a much more restricted use of 'phenomenology' to which I and most of the articles referenced, refer. "In recent philosophy of mind, the term "phenomenology" is often restricted to the characterization of sensory qualities of seeing, hearing, etc.: what it is like to have sensations of various kinds. However, our experience is normally much richer in content than mere sensation. Accordingly, in the phenomenological tradition, phenomenology is given a much wider range, addressing the meaning things have in our experience, notably, the significance of objects, events, tools, the flow of time, the self, and others, as these things arise and are experienced in our "life-world"."(Smith 2018).

Although Gallagher, in his works on direct perception, criticizes simulation theory often, he does not outright deny it has a role in social cognition. Gallagher argues (chiefly) against what he calls the 'strong pragmatic claim' for simulation theory (Gallagher 2001); the claim that what we most often experience in practice is best explained through simulation theory. He does permit simulation theory a *weak* pragmatic claim, that there are some rare examples of simulation being used in social cognition.

Instead, Gallagher proposes his direct perception as our most common means of intersubjective understanding. Direct perception purports that we are able to directly perceive other agents' mental states, without relying on any internal referencing or reasoning such as in simulating. "Perception is direct and smart. But it is also innately tuned to socially relevant aspects of the world. In contrast to this concept of direct social perception, on the TT and standard simulation theory accounts it seems that perception is only one part of a larger cognition process and that the social aspect is really found in what follows on from perceptual observation, in the application of folk psychology or in simulation routines, but it is not there in perception itself"(Gallagher 2008, p540). Gallagher first delves deeper into what it means to 'perceive', arguing against the view that perception is merely an 'input' to be processed through some internal inference.

This 'smart' perception is intended to better align with how we experience perceiving things than a 'dumb' input-only perception would. Gallagher argues that we do not experience taking in inputs and internally reason what these inputs must mean, but that we experience perceiving the meaning itself. E.G. when we see a car move by we simply experience seeing a car move by. We do not experience seeing certain shapes and colours, displacing over time. Subsequently reasoning that these shapes and colours match up with the idea of a car, and with its displacement past us over time, conclude a car is moving past.

The idea of perception being smart, as opposed to mere senses, is central to his theory. Interpreting perception in this way aligns with the phenomenological experience we usually have when understanding others. When we see happiness in another we experience directly perceiving said happiness, rather than experiencing first seeing certain expressions and internally reasoning that this must mean the other person is happy, which is made possible through this reinterpretation of what perception means. Although Gallagher admits there must be some neurological mechanisms involved in such a perception, such lower-level mechanical descriptions are not important when explaining this phenomenon in social cognition. Gallagher alternatively describes this idea of perception as a 'sensory-motor' process (Gallagher 2007), allowing us to perceive emotions and intentions in others, informing us directly about others, akin to a sensory process without internal reasoning or inference.

As such direct perception offers a second-person perspective on social cognition by allowing us to directly perceive the emotions and intentions others embody. This is opposed to the first-person perspective of processing mere sensory data internally, running it through a simulation, and coming up with conclusions about the other person from within our own point of view.

In later works, expanding upon direct perception, it has been proposed social cognition should be understood as broader interaction, with direct perception being a key part in *interaction theory* (Gallagher & Hutto 2008; Gallagher & Zahavi 2008; De Jaegher 2009). "I propose that the experiential directness of social perception can be understood only in the context of the role of *social interaction* in social cognition."(De Jaegher 2009, 535) Social cognition is understood as an interactive process. "According to IT, we understand the actions, responses, intentions and emotions of others, in their embodied comportments" ... "in the contexts of our dynamic interactions, all of which happen in the rich pragmatic and social situations of everyday life" (Gallagher 2015, 453). The meaning of words and social interpretation is altered based not on the one-sided perspectives of the different parties involved but is shaped through the shared interaction between them.

Now we understand a bit more about direct perception. Especially important is the direct access it provides to the intentions and emotions of others, as well as viewing direct perception as key to social understanding in light of a broader approach through interaction theory. Furthermore, it is important to note how Gallagher does not outright proclaim an incompatibility with simulation theory, but that he does describe it as particularly niche.

2.3 Apparent issues with combining these theories

Having briefly described simulation theory and direct perception in more detail, we already see hints at the many differences which could pose a problem for a combined theory. The simulation theorist's focus on explaining behaviour is in sharp contrast with Gallagher's(2001) later developed direct perception, which argues that *intentions* are what we primarily look at in social cognition. Gallagher argues for a second-person view of social cognition in which we (primarily) perceive what we understand of others directly. This opposes the first-person view of simulation, which describes understanding from the (simulated) perspective of the other. Furthermore, significant parts of Gallagher's descriptions of direct perception are written in contrast to, or as a critique of, simulation theory (Gallagher 2001; 2007; 2008). Even other authors stress the contention between the introspective and phenomenological approaches. De Jaegher(2009) interprets Gallagher's direct perception as part of a broader interaction theory, specifically as a means for direct perception to avoid being: "appropriated by the very cognitivist accounts criticised by Gallagher (theory theory and simulation theory"(De Jaegher 2009, 535).

How might these differences be overcome in order to find a combined theory which might be to the satisfaction of both perspectives? In order to better explore this question I delve deeper into the many conceptual differences between these approaches in the next section. I will discuss how we could approach combining these perspectives without letting these differences become incompatibilities. Furthermore, I will discuss specific issues Gallagher has with simulation theory and hybrids of it and direct perception, resulting in a perspective on combining these approaches which would preserve the unique perspectives these theories have to offer.

Armed with these important notes on simulation theory and direct perception, and the first inklings of the differences between them, we can start looking at how we should approach combining them.

3 Combining simulation theory and direct perception

3.1 Differences between approaches

There are many differences between how direct perception and simulation theory view social cognition. Before describing what the concept of a combined theory looks like I will first discuss key differences between simulation theory and direct perception, or between introspective and phenomenological approaches more generally. I will explain these differences and why they need not stand in the way of a theory which unites both.

Difference in explanandum

A distinction between introspective and phenomenological approaches to social cognition lies in the *explanandum*, that which is to be explained. Although both theories, in a broad sense, try to explain how social cognition works, there have been arguments about what this really means. Phenomenological accounts argue that mindreading accounts, such as simulation theory, lack a clear explanandum, and that phenomenology provides this. Phenomenological approaches try to explain the phenomena of social cognition; social cognition is understood as that which presents itself to us in our social experiences. "If we don't know the explanandum well, if we don't possess careful descriptions of its different facets - which is arguably one of the things that phenomenology can offer - it will, to put it mildly, be kind of hard to assess the relevance and explanatory power of the sub-personal mechanisms"(Zahavi 2011, 555). The potential problem for mindreading accounts is that because they do not aim to explain these experiences that they no longer have a clear answer as to what exactly it is that they are trying to explain. What is their explanandum if not what we experience in our social interactions?

Is the explanandum between both really that different? The problem, according to the phenomenologist point of view, lies with mindreading accounts describing social cognition primarily as explanation and prediction. This can also be seen in the earlier description of Gallagher describing social cognition (primarily) in terms of perceiving intentions. In this sense both sides of the playing field have a different understanding of social cognition, leading them to pursue an alternate explanandum. However, both still try to explain the same broader phenomenon - social cognition as a whole - merely moving the goalpost in an argumentative sense, rather than the overarching goal being different. Viewed through this lens it seems these authors are using a discussion on what the *explanandum* is as arguments for the value of their respective *explanans*, their explanation, instead. They argue about what 'social cognition' is by arguing which mechanisms are supposedly involved in social cognition or not - making it seem as if this were to exceed a discussion on explanans - rather than actually having a discussion on an entirely different topic of interest as you would expect from two theories which explain differing explananda. As the overarching explanandum social cognition, regardless of the precise mechanisms involved, remains the same. Such discussions should not make the goal of finding a combined theory which provides a broader explanans for this overarching explanandum any less feasible. For further reason to believe that the explananda of both aren't fundamentally different take Shannon Spaulding's (2018) stance arguing for a different terminology for mindreading theories. Spaulding argues that 'expla-

nation' and 'prediction' are not actually the right terms to describe mindreading accounts, but that 'interpretation' and 'anticipation' are a better fit. Phenomenologist authors agree that such terminology is much more in line with their understanding of what social cognition should be understood as, showing that the purported difference in explananda between the introspective and phenomenological approaches might only be a matter of perspective: "Under pressure, Spaulding (2018) admits that when characterising the subpersonal processes that allegedly underwrite social cognition 'one could substitute "interpretation" and "anticipation" for explanation and prediction' (p. 15). That is certainly closer to the mark and, if we are right, that move has much better prospects of bringing attempts at providing sub-personal explanations of social cognition into line with its actual character."(Hipolito et al. 2020) So it seems that despite having subtly different views on what is encompassed by 'social cognition', the overarching goal of these theories is much the same, and it should not stand in the way of a combined view. There might even be added value, by both sides offering a different perspective on the same explanandum, as a combination of the two which does not infringe on the separate perspectives, could provide an explanans for both.

Strong versus weak phenomenological constraints

How might these different perspectives on the explanandum influence their explanans? Despite the overarching explanandum of social cognition being the same, the difference in what these approaches tend to aim for does beckon more consideration. When your primary target is explaining an experience, as is the phenomenologist's aim, explanations for its underlying mechanisms risk receiving lesser attention. If, instead, your primary target would be to explain lower-level mechanisms, then your understanding of the experiential might still guide your explanation, but evidence on such underlying mechanisms is bound to receive greater attention. At first glance, this means that phenomenologist approaches are at an increased risk of under-explaining observations of lower-level phenomena. However, they would tend to better fit with our experience. On the other hand, a simulation theorist account might tend to have a more fleshed-out theory of the underlying processes, yet risk failing to adequately account for how their explanations fit the experiential level. These are, of course, a stereotype, but the aim of giving attention to, and explaining, phenomena by phenomenologist approaches has been criticized (Spaulding 2015). The problems identified in such criticism, I will argue, do not actually mean we should not strive to have our theories explain the phenomenological side of things.

Shannon Spaulding (2015; 2018) discusses these differences through what she refers to as the *weak* and the *strong* phenomenological constraints. The *weak* phenomenological constraint would have any theory on social cognition be at least *compatible* with our everyday experiences. The *strong* phenomenological constraint would have such theory not only be compatible but *explain* these experiences. In this sense, we can alternatively view the phenomenological perspective on the explanandum of social cognition, as described in the previous subsection, as an inclusion of the strong phenomenological constraint for theories on social cognition.

The point Spaulding aims to make, however, is that phenomenological theories imply that introspective mindreading theories, when rejecting phenomenological arguments, also reject the weak phenomenological constraint. She claims that phenomenological ac-

counts consider mindreading accounts incompatible with our experiences in social cognition. However, Spaulding argues this is not the case. When introspective theories merely reject the strong constraint; they do not aim to (fully) explain our experiences. "...one can accept that theories ought not make false claims about our phenomenology without accepting theories ought to explain and be empirically supported by our phenomenological experience." Spaulding argues that the weak constraint is necessary for any good theory on social cognition, as anything which contradicts our experiences clearly cannot explain social cognition correctly, our experiences are a part of it after all. Furthermore, Spaulding insists that the strong phenomenological constraint is not only unnecessary but unwanted; phenomenological evidence supposedly lacks epistemological value. "... phenomenological evidence ought to be novel, reliable, and relevant. However, phenomenological evidence lacks all three characteristics and therefore cannot play a substantial methodological role"(Spaulding 2015, 1088)

But are the shortcomings of phenomenology as evidence reason enough to ignore the strong phenomenological constraint? Spaulding has solid arguments against the epistemological value of phenomenological *evidence*, discrediting its reliability, and consequently, arguments based on such evidence. I, however, believe this does not fully negate the strong phenomenological constraint. Inconsistent results indeed make for poor evidence, but they are still very much a phenomenon which beckons explaining. In other words, the phenomenological evidence would serve a poor role as arguments for some explanans, but should still be considered part of the explanandum.

Even if one is to explain social cognition through underlying mechanisms, rather than the experiential level, then the experiential part of social cognition would still be a higher-level phenomenon relying on such explanations. Spaulding should agree with this, as she claims that any decent theory on social cognition requires being compatible with our experiences, as per the *weak* phenomenological constraint. As such the phenomenological level of social cognition could be said to supervene on the introspective mindreading accounts which Spaulding defends. As we know a supervenient phenomenon cannot change without the subvenient changing as well. In other words, our experiences of social cognition would not change without a change in the underlying mindreading mechanisms occurring as well (assuming those accounts indeed adequately explain social cognition on their own). As such, a complete understanding of how these subvenient mindreading mechanisms would work has a direct relation to the different expressions of the supervenient phenomenological evidence. So, if Spaulding's aim is to find a complete understanding of social cognition's underlying mechanisms through mindreading accounts, then this understanding too - with some translation - would necessarily explain the inconsistent and varied nature of the phenomenological level. This is why any account on social cognition which does not aim to (also) explain the phenomenological experiences, an account which only adheres to the weak phenomenological constraint, fails to fully explain social cognition. In fact, when phenomenological evidence is considered to be, in some way, part of social cognition, such as through adhering to the weak phenomenological constraint, then any full explanation of social cognition necessitates meeting the strong phenomenological constraint as well.

Simply put, a theory which fully explains social cognition cannot be compatible with

phenomenological evidence on social cognition whilst not being able to explain it as well. A thorough explanation of social cognition also needs to be able to explain these higher-level phenomena.

Why should we, more generally, actually strive to explain inconsistent evidence? Specific phenomenological evidence might not be novel, reliable, or (as evidence) relevant, however, this collection of experiences is very much a part of the social cognition that philosophy of mind aims to explain. The inconsistent nature of phenomenological evidence is instead an indicator that the phenomena involved in social cognition are rather complex and difficult to explain, and warrants further explanation. Take for example the magnetic properties of iron. Some samples of iron show that they are attracted to a broad range of different metals, whereas other samples show no such attraction. On the face of it, these pieces look the same. Can we now say that all iron is magnetic? Or perhaps that iron is not magnetic? Certainly not, for such extrapolation, the evidence is clearly inconsistent. However, leaving it at that means that we merely have evidence that we lack some understanding of iron's properties. These inconsistencies in evidence are an indication that the theory needs to be more complex and that the samples need further scrutiny, not less. As it turns out the atoms of iron can have their poles align. This can happen, for example, by being in close proximity to a magnet for some time, after which the previously non-magnetic piece of iron has become a magnet itself. The different samples of iron, with a more detailed explanation on the orientations of their atoms, are inconsistent evidence no longer, but consistent evidence for a more complex theory describing the properties of iron. If our experiences of social cognition are inconsistent there must be some relevant cause for which those inconsistent accounts are in fact consistent evidence. This means that the varied phenomenological accounts beckon further investigation, even though we should be careful not to base extrapolations on such inconsistencies.

Is Spaulding, as she claims, even really opposed to explaining inconsistent phenomenological evidence? Spaulding herself (2017, 2018) argues how judgements in social cognition can vary based on social categorization and in/out-grouping effects. Arguing for potential mechanisms for how people experience seemingly similar situations in social cognition differently. This strikes me as remarkably close to explaining exactly the sort of phenomenological evidence that she criticized before, and would be (partially) in line with the strong phenomenological constraint. Regardless, the inconsistency of phenomenological evidence is merely an indicator of the need for more complex theories which can encompass and explain those inconsistencies, rather than avoiding such evidence as a whole.

As such the issues with phenomenological evidence and the strong phenomenological constraint do not provide a conflict for integrating direct perception and simulation theory. Even more so it seems that there is actually good reason to strive to meet the strong phenomenological constraint. This should be done, however, with care as to argue for a theory using a small inconsistent subset of phenomenological evidence, but instead strive to explain the whole of it. Adding to simulation a theory better oriented for meeting this constraint could only help a combined theory get closer to solving the complex issue of the strong phenomenological constraint.

Cognitivism, mental representations, and the unobservability principle.

Another difference between these approaches lies with mental representation. Simulation theory represents other's mental states, within the simulations ran, in our own minds, conversely direct perception claims that we usually don't represent other's mental states at all, but we have direct access to them through our perception. This is closely related to 'cognitivism'; the idea of having internal reasoning as the basis for social cognition. Direct perception similarly objects to the need for such internal cognitivist processes for (most) social understanding.

A phenomenological purist might argue that there is no mental representation in social cognition whatsoever, which would be an issue for a theory which strives to combine a non-representationalist concept such as direct perception with a representationalist one such as simulation theory. Whilst Gallagher argues that the idea of representing mental states is not commonplace in social cognition he does permit simulation theory the weak pragmatic claim. The weak pragmatic claim being that simulation is something we use for the understanding of others in *some* (although in his words rare) circumstances (Gallagher 2001). Furthermore, Gallagher claims that the objective of simulation does not necessarily have to limit itself to mental states, permitting forms of simulation which could be considered non-representationalist (of mental states) to occur side-by-side with his broader interaction theory. "Moreover, even when theory or simulation is used, its target may be features of the situation or social context rather than mental states."(Gallagher 2015, 453) As such this divide should not necessarily pose a problem for combining these approaches.

What could, however, be a problem lies with the closely connected concept of the 'unobservability principle', the idea that thoughts and emotions are happening closed-off within the mind. "Mental states, and the minds that possess them, are necessarily unobservable constructs that must be inferred by observers rather than perceived directly" (Johnson 2000, 22) This concept is clearly opposed to Gallagher's concept of direct (social) perception, and Gallagher states that for any argument which seeks to unite direct perception with simulation theory, the unobservability principle has to be dismissed, or at least relaxed (Gallagher 2015, 453). On the other hand, introspective accounts such as simulation theory are classically associated with this unobservability principle, indeed an account of understanding others through internal simulation can easily be considered compatible with a lack of direct access to the minds of others.

How could an account of social cognition which combines both simulation theory and direct perception deal with the unobservability principle? First off, we do not need to look at the closed-off concept of thoughts as quite so black and white. I believe there are good reasons to believe that there are situations in which we do not have such access to other minds. Take a customer service representative handling another client's problem after a long day of work. They portray a happy and helpful character, both in their embodied comportments as well as throughout the dynamic interaction with the customer. On the inside, however, the service representative is exhausted, thinking only about the end of the workday and much-needed rest, portraying themselves professionally through rote. Does the customer have any kind of inference-free access to these thoughts and feelings of exhaustion? It does not seem

likely. Take however that same service representative with only their colleague around. Now they slouch and portray themselves exactly as exhausted as they feel. They are embodying their thoughts and feelings not merely because they no longer need to hide them, but also as a natural means of communicating those thoughts and feelings to their colleague. There would be good reason to interpret this situation as one in which the colleague does have access to the mind of the service representative.

An argument adhering to unobservability would be that in both cases the thoughts of the representative are hidden behind how they act. Although this may be true in some technical sense (i.e. there is in fact a body with expressions wrapped around the mind experiencing these thoughts in between the mind and observer), it seems an unfair representation of what is experienced in such scenarios. In the first the service representative is hiding their true thoughts, perhaps subconsciously, but in the second they are merely letting loose. As such, it seems more fair to say that the accurate embodiment of these thoughts and feelings would be the default, but that there is some ability to alter said embodiment. When we stub our toes we are naturally inclined to wince, perhaps even exclaim our pain, but we are also capable of suppressing this; perhaps we stubbed our toe next to a sleeping baby and are trying not to wake them. Such a view would permit that we are capable of directly perceiving other people's mental states to some extent, but that we are also capable of using our embodied comportments not as a natural channel into our minds, but as a means of hiding them. Furthermore, I take it likely that there are many thoughts and feelings we have which are too complex to understand even through honest embodiment alone, as we will see an example of when I sketch out the idea of a combined theory in the following subsection.

When I look at the desk I am writing at I perceive wood, even if I technically only see the light reflecting on the wood as it has been altered by the clear oil covering it. When I look at my cabinet down the hall, however, I do not perceive wood, despite it being made of pine, as it is fully covered by a smooth non-translucent paint. Despite some technical lack of direct observability when looking at both cases, the former experience would be more accurately described as being able to directly perceive the wood of my desk. Overall the notion of unobservability seems appealing for certain situations, and perhaps true for all things 'mind' in some technical sense. But often what we experience would be more fairly described as having direct access to other minds, particularly when we are openly and honestly communicating bodily, undiminished by such bodily communication technically being all we can see. As such I see good reason *not* to strictly adhere to unobservability, and instead look to combine the introspective with theories which describe direct access, such as direct perception.

We have looked at the relevant theories and seen how these approaches differ and relate, and seen a glimpse of the explanatory advantages a combination might offer. I will now discuss arguments Gallagher levies against simulation theory and see how this might apply differently to a combined approach, to further judge the compatibility of the two. I will also discuss what sets this idea of combining the introspective with the phenomenological apart from previous hybrid theories, theories which Gallagher criticizes. After this, we will understand the potential merits of a combined approach and I will give an example of what such a combined theory, which seems to walk the fine line between being pluralist or hybrid, might look like.

3.2 Gallagher against simulation theory

The strong pragmatic claim

As mentioned in section 2.2 Gallagher(2001) argues chiefly against the *'strong pragmatic claim'* for simulation theory (and theory theory). The strong pragmatic claim for simulation theory would be the claim that the primary way in which we understand others would be through simulation. The *weak* pragmatic claim would merely claim simulation as a tool for social understanding in less common situations. The weak claim is left uncontested, but for the strong pragmatic claim Gallagher offers his alternative: "The alternative that I will propose is that the understanding of the other person is primarily neither theoretical nor an internal simulation, but is a form of embodied practice."(Gallagher 2001, 85) Important to note is that, despite Gallagher not contending the *weak* pragmatic claim, he does describe the cases of social understanding in which methods such as simulation theory or theory theory might be employed as exceptionally rare in comparison. This is to say that Gallagher's view does not align with that of both direct perception and simulation theory being heavily and readily used tools for understanding, with direct perception merely taking the favourite spot. Instead, in his view, mindreading accounts such as simulation theory would be lying in a drawer somewhere, to be taken out only for niche applications. As such, Gallagher arguing against the *strong* pragmatic claim of simulation theory is still very much against my proposed view of a more intermingled use of both tools.

For backing up the strong pragmatic claim for direct perception Gallagher claims we usually interpret actions on an intentional level: "For example, if, as I reach for a cup, someone asks me what I am doing, I do not say, ordinarily, 'I am reaching for a cup': rather I say, 'I'm taking a drink'. I tend to understand my actions just at that pragmatic, intentional (goal-oriented) level, ignoring possible sub-personal or lower-level descriptions, and also ignoring ideational or mentalistic interpretations, e.g. 'What are you doing?' 'I'm acting on a belief (desire) that I am thirsty'. Likewise, the interpretation of the actions of others occurs at that same pragmatic (intentional) level."(Gallagher 2001, 88) His further argument is that we are able to directly perceive intentions in others as they are embodied by them, rather than needing to rely on mindreading. "In effect, this kind of perception-based understanding is a form of body-reading rather than mind-reading. In seeing the actions and expressive movements of the other person one already sees their meaning: no inference to a hidden set of mental states (beliefs, desires, etc.) is necessary."..."Importantly, the perception of emotion in the movement of others is a perception of an embodied comportment, rather than a theory or simulation of an emotional state."(Gallagher 2001, 90)

Gallagher's argument against the strong pragmatic claim for simulation theory in these sections can be summarized as follows: 1: Gallagher claims that our main means of intersubjective understanding is the reading of *intentions*, rather than (possibly associated) beliefs/desires/mental states. 2: Gallagher claims that we are able to directly perceive intentions as they are embodied in the movements and expressions of others. As such there is no need for simulation as our primary means of intersubjective understanding. To make my view of an intermingled theory itself have a strong pragmatic claim I would have to show that the

proposed direct perception is not so prevalent and/or that simulation might itself be used more often than Gallagher thinks.

Why direct perception might not be as prevalent as suggested

Now take a converse example. Imagine a police officer during a random car stop asking a citizen for their licence. The citizen casually reaches for their back pocket to grab their wallet which holds the document. In this situation, the police officer might accurately perceive the intent of the citizen being one of compliance. Imagine instead that the officer is tremendously racist against people of the citizen's descent. The exact same scenario happens, however in this scenario the racist officer inaccurately perceives the citizen to have a hostile intent; the officer is convinced the citizen intends to reach for a weapon. In neither scenario does the officer *see*[4] the intent of the citizen; the officer cannot see what the citizen is grabbing. In the first scenario, the officer would conclude that the citizen is grabbing their license and patiently waits a moment. However, the racist officer, fueled by the (inaccurately) perceived hostile intent, believes the citizen to mean him harm, leading the racist officer to respond with swift violence. In the two versions of this example, the embodied intention of the citizen remains entirely the same; one of compliance. The *perceived* intention however changed because of the officer's bias.

Shannon Spaulding offers a possible understanding of how such bias might lead individuals to different conclusions in social cognition: "*We often find ourselves offering divergent interpretations of others' motives, intentions, beliefs, and emotions.*" ... "*I argue these disparities in social interpretation stem, in large part, from the effect of social categorization and our goals in social interactions*"(Spaulding 2017, 409). Following Spaulding's view the social categories to which both the subject and observer belong often influence their judgements in social cognition. The racist officer from the example might belong to some political persuasion which holds people from the ethnicity of the citizen in ill regard, influencing his judgement as such. *"Social categorization influences what we attend to and how we interpret social interactions. Categorizing individuals, behaviours, and events depends on the situational context, our experiences, and our expectations, and these may differ predictably and dramatically among individuals with different social backgrounds. Augmenting existing accounts of mindreading with these data will allow us to explain how social disagreements arise and provide us with a more realist, accurate account of how we understand others."*(Spaulding 2018, 426) It could be argued that these social categories are identified based on what people embody; the racist officer identifies the citizen's ethnicity based on appearance alone, however, this does not make the bias caused by these categories part of some embodied social cognition as per direct perception. The officer's bias against the citizen's racial category relies on internalized knowledge about what that category must mean about their intentions. The altered perception does not come from what the citizen embodies but from the officer's own mind. Certainly, such bias is a part of social cognition which cannot be understood through direct perception, and we know that bias of all sorts influences our judgements, meaning that this also could form a problem for direct perception's strong pragmatic claim.

[4]Here I use seeing versus perceiving much in line with Gallagher as discussed in the introduction. What Gallagher refers to as perception being what we experience perceiving and 'smart', versus the 'seeing' being raw sensory data.

For further problems with direct perception of intentions let's take a look at another example. Again a citizen is stopped for a random check by a police officer. The officer again asks for his licence. The citizen, however, has an expired driver's license and intends to deceive the officer by passing it off as legitimate, to do this the citizen puts on a convincing act of a regular friendly citizen handing over their licence, so as to not arouse suspicion. The police officer interprets the citizen's intent as casual compliance and scans the document without suspicion, missing subtle flaws which might have alerted him to the document's illegitimate nature. In this example, the officer perceives what the citizen wants him to perceive. It can be argued that the officer in fact accurately perceives that which is embodied, but the citizen intentionally does not embody the deceitful nature of his actions, so even interpreted as such, the officer fails to perceive the citizen's intent. Does this mean that deceitfully embodied intent is simply something out of the norm, something which indeed cannot be directly perceived, but which might also be rare enough to protect Gallagher's strong pragmatic claim? Take a tweaked version of this scenario. Everything starts out the same, however, in this version, the officer recognizes the citizen as a local crook. This crook does not have a warrant out on him as far as the officer is aware, but this recognition is enough to make the officer extra alert. The officer interprets the crook's intents not as casual compliance, but as some sort of deception. The police officer now examines the documents thoroughly and spots that the license is in fact expired. In this last scenario, the officer correctly interprets the intent of the crook, however, he does so by partially disregarding what the crook embodies, basing the intention he perceives in part on previous knowledge.

With these examples we have identified three scenarios in which direct perception fails to explain the perceived intentions. 1: Perception has been altered based on bias, inaccurately concluding intentions which are not embodied. 2: The embodied intent is an act of some sort, and the acted-out intention is perceived rather than the actual intention. 3: The embodied intent is an act of some sort, but knowledge of deception bases the interpreted intention not (just) on the embodied intent, but also on some imagined deceit. As such, the strong pragmatic claim for direct perception, that direct perception would be by far the primary way through which we understand others seems like an exaggeration. Scenario's like these, where our perception of intentions is altered based in part on something more than what we can directly perceive, seem quite common indeed, and such scenarios only describe the intention-reading part of social cognition being done through other means than just direct perception, going against Gallagher's(2001) strong pragmatic claim being based on direct perception explaining intention reading. Examples such as Gordon's(1992) chess player trying to predict the opponent's moves by simulating their perspective, are types of social cognition which go beyond the reading of intentions or plainly visible emotions, and direct perception does not even claim to be able to explain those parts. Overall this leaves direct perception's strong pragmatic claim appearing rather weak indeed. Although direct perception might still be used often, as very straightforward intention-reading is a regular part of social cognition, there seems ample reason to think that other forms of social cognition also happen regularly, and as such, direct perception cannot hold the strong pragmatic claim on its own.

Does simulation provide understanding of the other?

Gallagher argues that simulation fails to provide an actual understanding of the other:
'If I project the results of my own simulation onto the other, I understand only myself in that other's situation, but I don't understand the other. In other words, given the large diversity of motives, beliefs, desires, and behaviours in the world, it is not clear how a simulation process based on my own relatively narrow experience (or relatively unique circumstances) can give me a reliable sense of what is going on in the other person's mind, or in their behavior.' (Gallagher 2007)

Is this a fair criticism of simulation? This argument Gallagher posits against explicit simulation appears to attack a straw man of said theory. Gallagher here interprets simulation fundamentally differently from what Gordon and Goldman describe. Goldman describes simulation to involve creating a context that is suited to the target, feeding their beliefs, behaviour, history, etc. into the simulation before using said simulation to draw conclusions on their mental state (Goldman 2006). This means that simulation theory does account for the diversity of motives, desires, etc. that others have, and does not reason from the observer's point of view. According to this description, although simulation might occur within the mind of the observer (just as direct perception necessarily would), it happens by substituting one's own context for the other; it simulates the other, providing *only* understanding of their perspective. In the same vein, Gordon's idea of a partial projection(Gordon 1992) is simulating what the situation would be for the other, substituting for their context. Only the low-effort total projection variant would line up straight-forwardly with Gallagher's critique. These simulation processes do not simulate based on a person's 'own relatively narrow experience (or relatively unique circumstances)', in fact, they are described as substituting these for that of the other. As such, Gallagher's counterargument here appears to attack a straw man of simulation theory, or at least is an adversarial interpretation of simulation as merely the description of 'putting oneself in another's shoes', which Gordon (Gordon 1992) has argued against, as described in section 2. Instead, the varying degrees of projection allow for varying degrees of understanding the other's perspective moreso as the self through total projection, or moreso as the other when projecting less and less context of the self onto the other, through more intense simulation. Simulation theory as described by these authors is not something people use to answer the question 'What would I think and do in their shoes?' but to answer 'What do they think and will they do?'. As such, the argument that simulation does not provide understanding of the other is flawed.

However, a more well-willing reading of this passage by Gallagher might lead one to believe that he instead intended to argue that there is no access to the *actual* beliefs and desires of others because simulation is taking place introspectively. This could mean that simulation which *would* have such access might still understand others to Gallagher's satisfaction, but in practice there would be only some sort of faux-access; one can merely substitute the context and beliefs of the other as they see it. This faux-access would result in the simulation being indirectly based on one's own (too) narrow set of experiences rather than those of the other. This argument might seem very similar to the previous interpretation, but it is important to be precise about these things if we wish to fruitfully discuss them, as an important distinction has been made. The first (and given) argument claims that the theoretical model

of simulation itself is flawed, but they instead attack a straw man of said model. What I take to be the intended, and more troublesome, argument states that the given simulation theory model, whilst sound within its own scope, requires knowing the context of the other, which one could argue that simulation does not provide access to.

You might now think my counterargument merely points out an unfortunate phrasing by Gallagher, which, when worded more to my satisfaction would still prohibit simulation from providing the required understanding of others. Indeed if we instead view the argument as one against the access that simulation theory would have to the prerequisite context of the other the argument seems solid. The point I am trying to make, however, is to specify that the valid part of Gallagher's criticism doesn't attack the concept of simulation itself but merely brings into question the access that simulation has to the other's context, which is a prerequisite for simulation to provide the needed understanding. With access to the context of the other person, rather than some indirect guess at it, simulation theory would still lead to simulating the other person's perspective, rather than the self in it. As such, providing understanding not of the self in the other's shoes, but of the other. In this light, Gallagher's argument might pose a problem for a poorly-informed simulation theory, with direct perception as an alternative with better-explained access to the thoughts and feelings of the other. However, this well-willing view of Gallagher's criticism simultaneously elucidates a complementary role that direct perception could have for simulation within a combined theory. Direct perception could provide simulation theory with the kind of access to the context of the other it needs to understand them.

3.3 Combining the theories

How should we approach combining these perspectives? The most straightforward way to combine simulation theory and direct perception would be to take implicit simulation theory, which does not postulate an experiential level, and claim it as the backbone of direct perception. This seems especially amenable given Gallagher does accept the necessary presence of unspecified lower level mechanisms to direct perception: "Perception does involve complex processes at a sub-personal or unconscious level, but this does not make perception, which is an activity of the perceiving subject (organism), indirect..."(Gallagher 2008, 537). This way of integrating both theories, however, is exactly what De Jaegher considered a "...hijack into an inferencing/simulating account in which social perception builds on fast and behind-the-scenes cognitive processing."(De Jaegher 2009, 535). De Jaegher aims to prevent such appropriation by extending direct perception to be inseparably viewed from interaction.

I do not subscribe to the notion that social cognition should only be viewed as interaction nor merely embodied; there seem to be parts of social cognition which *are* indirect, such as understanding a character in a show or trying to understand the thoughts somebody is hiding, as well as cases of partially indirect social cognition, such as when you are participating in some social activity but your understanding continues to form independently even when there is a pause in interaction. I do, however, find the notion that interaction is an important part of social cognition congenial. A phenomenologically inspired view on direct perception, rather than being a mere experiential layer on top of implicit simulation such as

the integrative account De Jaegher fears would have it, seems better aligned for understanding such interactions. Correspondingly I will not aim to explain such a superficial combination(hijacking) of the theories. Rather I intend to subscribe to the notion of direct perception as it stands, which offers value for understanding interactive social sense-making, and is compatible with Gallagher's broader interaction theory (Gallagher & Hutto 2008; Gallagher & Zahavi 2012; Gallagher 2020). What I am looking at is to judge the compatibility of an unadulterated direct perception with simulation theory, offering explanations for both the embodied and enacted parts of social cognition as well as those parts which are less direct.

How does this approach to combining the theories compare to previous attempts? Conveniently Gallagher has discussed previous attempts to integrate theory theory or simulation theory with direct perception. Gallagher argues against these hybrid theories, in favour of a pluralist account. Gallagher describes how what he refers to as hybrid theories aim to integrate direct perception to fit the perspective of the introspective theories, threatening direct perception's phenomenological perspective. A pluralist account, in Gallagher's view, would instead allow for varying theories on social cognition to be accepted as separate entities accounting for social cognition. "... I'll suggest that a truly pluralist approach that recognizes many contributories to a full account of social cognition, rather than a new hybrid approach, is a better alternative."(Gallagher 2015; 453). Gallagher argues that hybrid theories (Bohl & Gangophadyay 2013; Michael et. al. 2014), which co-opt the inner workings of both theories, either weaken the concept of mindreading or fail to properly accommodate direct perceptual understanding. He concludes that a pluralist approach which treats the various theories as separate strategies is preferable to the compromising middle ground such hybrid theories provide.

The sort of combination we are exploring, however, aims to maintain both unique perspectives of simulation theory and direct perception, rather than having a hijacked view of either process. In this sense, the combined theory I sketch out might be more fairly characterized as pluralist. The difference between approaches becomes even more evident when looking back at the discussion on De Jaegher arguing about the dangers of direct perception being hijacked by cognitivist accounts on social cognition(De Jaegher 2009). This I sidestepped by not resolving a combination of theories through adopting the sub-personal mechanisms of direct perception as some form of implicit simulation, as De Jaegher feared, but instead leaving the phenomenological take on direct perception intact, which is needed for it to be able to offer the explanatory advantages of that perspective. In this sense my look at the compatibility of these theories is not the same as the more straightforward attempts at integration done to create the hybrid theories which Gallagher warns about, but much closer to the exploration of a pluralist account for which he advocates. That being said neither is this combined view as simple as the pluralist idea of separate strategies which can coexist; I suggest these strategies for social cognition could very well be intermingled and complementary, which requires further thorough research at precisely how these processes can inform each other and when they are active. As such, despite the critical look at various arguments by Gallagher throughout this thesis, this exploration is very much a step in the direction of working out a pluralist account of these processes for which he too advocates.

A combined view

How might a combined view which both leaves independent theories intact, whilst also letting them inform and complement each other, look like? Direct perception could be viewed as capable of providing context for simulation and vice versa. Let me illustrate the concept by showing how social cognition might make use of both simulation theory and direct perception in this intermingled way through some examples.

Imagine a person telling an acquaintance a crass joke, expecting beforehand that the joke would be appreciated. However, an irritated expression flashes across the acquaintance's face. The jokester immediately perceives that the joke didn't land as intended and their acquaintance is upset. It stands to reason that this can be directly perceived and understood through the irritated expression seen, rather than needing some simulation process. The emotions they are experiencing are clearly visibly embodied through this irritated expression. What the jokester, however, does not yet understand is exactly *why* they are irritated. Perhaps the acquaintance took offence at a stereotype the joke relied upon. Or could they be irritated for an entirely unrelated reason?

The irritation can be directly perceived, but the specifics behind the irritation cannot. There might be clues, in the timing, and broader embodiment, but the possibly intricate reasons would be far too complex to be understood through the perception of the irritation alone. Due to the potential awkwardness or soured relations, if the missed joke indeed caused the irritation, it is pertinent for the jokester to better understand the reason behind it.

Amending this direct social perception with related ideas from the broader interaction theory might have you think something like the following would happen: The jokester directly perceives the irritation. The jokester also understands the need for a better understanding of said irritation. He is moved to ask 'I'm sorry, did I offend you in some way?' The acquaintance might respond: 'I've never understood that stereotype, having lived with [stereotyped group] for a long time it has always struck me as untrue and unfair. As such, I do not appreciate the joke.' Now the jokester, and the acquaintance both, have a deeper understanding of the previously perceived irritation and know how to act accordingly.

This view, however, seems to miss out on some important steps of what would happen. When the jokester perceives the irritation alarm bells start ringing. Better understanding is needed, and even a slightly more attentive perception of the embodied does not reveal this meaning. The jokester might understand that this needs a question for clarification; revealing the meaning through further interactions, however, the jokester will most certainly immediately try to figure it out himself. Even before the jokester has finished asking after the offence, he would be wracking his brain to find a better understanding himself. Taking the caused irritation as a new piece of context used to inform simulation might even let them realize that the acquaintance not only lived with the stereotyped group but apparently takes offence at the stereotype because of it. Or perhaps the jokester, even with the amended context he has available through direct perception still knows too little to run a successful simulation and remains in the dark until the acquaintance elucidates him. However, the need for further understanding surely moves the jokester to attempt this through inference as well, rather than merely waiting on some pivotal response. Imagine the acquaintance re-

fusing to elaborate and walking off, would the jokester plainly accept that there is some lack of understanding that he will never perceive? Or would the jokester try to solve his confusion by amending perception with a simulation-based inference? In a scenario that is pressing enough the jokester might even simulate it consciously, going as far as to experience the process of trying to fill in the gap in understanding through repeated simulations. However, even in the split second after perceiving the acquaintance's irritation, the context through which the jokester understands him will be altered and rerun, with a newfound understanding dawning on the jokester even before the reason would be explicitly stated.

Imagine instead that the acquaintance attempts to hide their irritation, and they politely laugh at the joke. The jokester, who still believed the joke to land well, would no longer perceive irritation, but instead might (incorrectly) perceive enjoyment. Now both the initial implicit simulation of the acquaintance, as well as what is perceived of them, is wrong. In this scenario, later in the evening, the jokester overhears somebody else telling a story referencing how the acquaintance used to live with the stereotyped group from the joke. The jokester might suddenly realise that his implicit simulation; his initial idea that his joke will land well, was missing out on some important context. In fact, it must have been wrong. The jokester substitutes this newly realised variable and now imagines how the acquaintance would actually feel about the joke, and how he might not be inclined to let it show. The jokester remembers how the joke landed, but in his mind's eye he no longer perceives the polite laugh as enjoyment, but rather as the hidden irritation it truly represented.

What is the role of simulation in these examples? We could view the simulation used throughout this perception as having various degrees of the previously described 'projection'(Gordon 1992). We might imagine an unachievable 'ideal' simulation as having no projection at all; substituting every single piece of context of the other person in place of our own. On the other end of the simulation-spectrum would be Gordon's concept of total projection, which does not substitute context. Initially the jokester, before deciding to tell his joke, would implicitly simulate his acquaintance with total projection. He would wrongly reach the conclusion that the acquaintance would surely enjoy this joke. Immediately after perceiving the irritation the jokester realises he was mistaken, but not yet exactly how. Now he will run a more effective *partial* projection, using up greater resources and substituting at least some minimal context; that the acquaintance has a reason for being irritated. If the jokester wasn't overly alarmed by the initial irritation this second simulation might still substitute little; a minimum additional effort in order to amend his conclusion, and this might still happen subconsciously. If, however, the jokester was greatly disturbed at the prospect of a social faux pas, he might try to make this partial-projection project as successful as possible. He would then try to substitute every piece of context that he can remember, identify, or *perceive* of the acquaintance. This kind of simulation would be closer to the ideal of simulation without projection, yet depending on if the jokester's knowledge (and perception) is (in)accurate, might still substitute some context wrongly. Such simulation would be much more resource intensive, lining up with what the jokester would experience if he panics and *consciously* wracks his brain to understand the acquaintance's response more clearly. Then we see where simulation crosses from being implicit to being explicit, from a subconscious

process to a conscious one.

Explicit simulation is fairly described by Gallagher as a means of social cognition which, though it might well occur, certainly isn't our primary means of social understanding (Gallagher 2008). Most of the time we would understand others clearly enough through direct perception or the somewhat more intensive implicit simulation, rather than needing such active replacement of context to prevent such as the panicked jokester's explicit simulation would require.

Instead here occur various instances of total and partial-projection which implicitly simulate the other in an attempt to understand their mental states to a specific degree. First some quick total-projection implicit simulation before telling the joke, which leads the jokester to expect it to be appreciated. Second a slightly more intensive partial projection before a continued lack of understanding leads to panic setting in, followed finally by explicit simulation which attempts to project as little as possible from the perspective of the jokester. In the meantime, the other's intentions and emotional states are being directly perceived. In the earlier example, the initial implicit simulation leads to a wrong conclusion, and direct perception of the acquaintance's irritation is what makes the jokester realise this. That leads him to amend his understanding both through more extensive simulation as well as through interaction. In the example where the acquaintance politely hides his irritation, however, both the initial simulation and direct perception turned out to be incorrect, and it is a later simulation which leads to both better understanding and an altered perception of that specific laugh of the acquaintance. As such, this view intermingles both the low-level implicit and high-level explicit simulation with direct perception, and understanding it in this way shows not only how they both are at times used, but how they are used to amend and supplement each other.

These examples illustrate an intermingled use of direct perception and simulation, where both can complement each other. We can imagine many degrees of simulation occurring side-by-side to direct perception. Explicit simulation, which Gallagher argues is very rare, indeed only occurs during the jokester's panic. The rarity of this resource-intensive explicit simulation, however, does not stand in the way of simulations of varying resource intensity being used in order to try to understand the acquaintance more introspective means, complementing the direct access to intentions and (surface) emotions which direct perception provides.

On the other hand there is reason to be wary of infringing upon either theory's perspective. As seen Gallagher or Spaulding might argue that a combined theory, by allowing the interaction and complementary informing between these approaches, could encroach on the phenomenological reasoning behind direct perception and infringe its perspective as the before mentioned 'hybrid' theories do. However, no attempt is made to hijack the underlying mechanisms of direct perception, nor a strict adherence to the unobservability principle is shown to be necessary for combining these theories. As such, Spaulding's concerns are sidestepped, and the proposed ideas in this thesis appear more in line with Gallagher's desire to work at a pluralist understanding of social cognition.

Another potential issue might then be found through the opposite perspective; from the hybrid theories which aim to integrate the introspective and phenomenological by adopting direct perception into some introspective mindreading account. Such a hybrid theory might argue that a combined theory which refuses to account for direct perception in some introspective manner also fails to adhere to the unobservability principle. To such criticism I have to give *some* ground; I do not believe that a combined view which respects the unique perspective of direct perception and the broader theories surrounding it can be reconciled with strict adherence to unobservability. If one truly believes other minds to be inaccessible except through inference, this would have to infringe upon the essence of direct perception. Any hybrid theory which follows strict adherence to the unobservability principle - by adopting direct perceptions to be mechanically understood as introspective in some way - would, however, also fail to retain the unique phenomenological explanatory advantages which direct perception provides. Such a theory might be a combined theory in some sense, but it would fail to augment its explanation with the perspectives that my approach to a combined theory strives to preserve. By sacrificing this phenomenological perspective, such hybrid theories would fail to meet the strong phenomenological constraint as much as any singular introspective theory. Accordingly, those overly integrative hybrid theories should not be considered a combined theory of the same sort as which this thesis is taking a closer look at. Rather such hybrids would reside solely in either the introspective or phenomenological side of the debate, instead of offering an approach which combines these perspectives.

We now have an amiable look at what a theory which combines simulation theory and direct perception might look like. They could both co-exist, providing explanations for those parts of social cognition for which they offer the best understanding. We also see how they could complement each other, providing a new understanding of those social scenarios in which there does not seem to be a clear answer to whether we would only use either theory. When having looked at the many differences between these accounts, as well as the differences between the overarching introspective versus phenomenological inspiration, it is clear that most of this conflict does not constitute incompatibility. A combined theory might be made which retains both unique perspectives, yet allows them to interact enough to complement each other. Certainly, there is ample reason to further explore more specific fleshed-out combined theories which can straddle this line between being pluralist, without losing the complementary interaction, and hybrid, without compromising unique perspectives.

4 Conclusion

We have seen how the introspective and phenomenological perspectives have some differences in explananda, and how a theory which *does* strive to adhere to the strong phenomenological constraint might be able to meet these added criteria the phenomenological perspective has for their view on the explanandum of social cognition. Furthermore, we now know that any theory which aims to fully explain social cognition cannot do so without meeting the strong phenomenological constraint, which provides good reason to amend introspective theories with phenomenological ones better suited to this task.

On the other hand, direct perception has been shown to have trouble holding the strong *pragmatic* claim on its own, failing to address a variety of situations which occur in social cognition such as those involving bias or deceit, or those involving situations where we lack direct access to the inner thoughts of others. Making it seem more feasible for simulation theory and direct perception to hold onto the strong pragmatic claim when united. Furthermore, we have seen how *not* to approach combining theories, by integrating the phenomenological perspective of direct perception into an introspective account, which would fail to offer the unique explanatory advantages provides. Finally, we saw a first look at how a combined theory might look which maintains these explanatory advantages, by respecting both the introspective and phenomenological perspectives that simulation theory and direct perception provide.

How does this view on uniting these theories on social cognition compare to other attempts at combining theories in this field? We have discussed the differences between phenomenological and introspective approaches to social cognition, in particular with respect to direct perception and simulation theory. I have already shown many reasons why these approaches might be more compatible than these differences would make them seem, and that a combined theory might be able to grasp both of the otherwise unique perspectives and advantages of these different theories. This, however, is not the only attempt at combining various theories on social cognition, although such a direct combination of the introspective and phenomenological approaches is uncommon, there have been more temperate attempts at this as was briefly mentioned with for example Goldman's moderate approach to embodied cognition (Goldman 2012). More commonly, however, we see hybrids within the various introspective theories or within the various phenomenological theories. There is, for example, a lot of work done in philosophy of mind on combining theory theory and simulation theory (Currie & Ravenscroft 2002; Nichols & Stich 2003; Goldman 2006). On the other side, we regularly see direct perception integrated into various takes on participatory sense-making or interaction theory (Gallagher & Hutto 2008; Gallagher & Zahavi 2008; De Jaegher 2009), as well as the overarching tradition of 4E cognition itself looking at these various aspects of social cognition as a whole.

This thesis does not delve deeper into such hybrid theories, nor does it really compare itself against them. This is not, however, because I wish to discount them, on the contrary; while this theory looks to bridge the gap between the introspective and the phenomenological it does not have to contend with those theories which more firmly ground either side into a more coherent whole. I believe social cognition to be a very broad and complex part of our lives, both experientially and mechanically. While I find simulation theory and direct

perception to be particularly attractive theories, which explain much of social cognition, *especially* when combined, I think there is a lot to be said for amending these theories further. There are many more tools on our social cognition toolbelt, and I see no immediate conflict with these other combined theories from either the introspective or the phenomenological domain. Instead, I see the potential for expanding and integrating our theoretical explanations even further, to better represent all the tools social cognition uses.

The intermingled use of the theories explored in this thesis could also be taken as a step towards not only combining direct perception with simulation theory, but as a step towards combining phenomenological theories and introspective theories as a whole. An exploration of such more generalised combinations is, in my opinion, very much warranted. While this thesis does provide some steps in the direction of such exploration its focus is very much on simulation and direct perception. Firstly because of the body of work of criticism between them. And secondly because, as we've seen, direct perception might be employed in social cognition to provide simulation with the kind of direct access which a purely introspective account of simulation theory would lack. As such, the combination of these theories, in particular, seems interesting for just such a combination which maintains the integrity of both, as a pluralist account would, as well as allow for a complementary interaction between them, which veers somewhat towards the advantages a more integrated hybrid theory might provide. I find this complementary action between theories especially valuable as such action provides an avenue of tangible value a combined theory might provide over simply using the theories separately, which a more segregated straight-forward pluralist account would struggle with.

Throughout this thesis we have seen that despite the many differences between these theories there is reason to believe they are not insurmountable. Having looked at differences and conflict both in an overarching sense between the introspective and the phenomenological approaches, as well as more specifically looking at criticism particular to simulation theory and direct perception, there are many reasons to think these theories are compatible. Even more so, a combined theory might better explain social cognition as it is able to retain advantages unique to both perspectives.

Could a combination of simulation theory and direct perception lead to a theory which better explains social cognition than its parts? I would argue, yes. A lot of the conflict between these theories boils down to their perspectives arguing that they explain *more* of social cognition and that they are able to better capture especially relevant aspects of it. Such conflict, however, might merely be considered arguments for why their parts of the explanation might be the most important but is not an argument against combining them. In fact, it appears that those same arguments provide reasons to believe we should be looking at combining the two, as it could provide a theory which explains all these different aspects of social cognition, which were argued to only fit their unique perspectives. We now also know that these theories might be combined in a way which respects those unique perspectives, rather than infringing upon them by 'hijacking' one theory with the other, or finding some compromising middle ground. Instead, they can coexist in a theoretical understanding both as separate

tools with their unique uses as well as informing each other, providing a more comprehensive pluralist understanding of social cognition, but one which does not shy away from viewing the theories it contains as something which interact and complement each other.

Further exploration of the inner working of how such an intermingled view of simulation theory and direct perception is needed. Particularly such exploration which allows simulation theory to remain compatible with other introspective accounts on social cognition, and similarly, having an account of direct perception that remains compatible with its integration into a broader interaction theory.

5 References

Bohl, Vivian and Gangopadhyay, Nivedita. 2013. "Theory of mind and the unobservability of other minds" *Philosophical Explorations* 17(2), 203–222.

Currie, Gregory and Ravenscroft, Ian. 2002 *Recreative Minds: Imagination in Philosophy and Psychology* Oxford: Oxford University Press.

De Jaegher, Hanne. 2009. "Social understanding through direct perception? Yes, by interacting" *Consciousness and Cognition* 18:535-542.

Dennett, Daniel C. 1985. *Brainstorms* Cambridge, MA: MIT Press.

Gallagher, Shaun. 2001. "The Practice of Mind: Theory, Simulation or Primary Interaction?" *Journal of Consciousness Studies* 8:83-108.

Gallagher, Shaun. 2007. "Simulation trouble" *Social Neuroscience* 2:353-365.

Gallagher, Shaun. 2008. "Direct perception in the intersubjective context" *Consciousness and Cognition* 17:535-543.

Gallagher, Shaun. 2015. "The new hybrids: Continuing debates on social perception" *Consciousness and Cognition* 36:452-465.

Gallagher, Shaun. 2017. *Enactivist Interventions* Oxford: Oxford university press.

Gallagher, Shaun. 2020. *Action and Interaction* Oxford: Oxford university press.

Gallagher, Shaun and Hutto, Daniel. (2008). "Primary interaction and narrative practice" *The shared mind: Perspectives on intersubjectivity* x:17–38

Gallagher, Shaun and Zahavi, Dan. (2012). *The phenomenological mind* Routledge

Gallese, Vittorio and Goldman, Alvin. 1998. "Mirror neurons and the simulation theory of mind-reading" *Trends in Cognitive Sciences* 12:493-501.

Goldman, Alvin. 1989. "Interpretation Psychologized" *Mind & Language* 4:161-185.

Goldman, Alvin. 1992. "In Defense of the Simulation Theory" *Mind & Language* 7:104-119.

Goldman, Alvin. 2006. *Simulating Minds* Oxford: Oxford university press.

Goldman, Alvin. 2012. "A Moderate Approach to Embodied Cognitive Science" *Rev.Phil.Psych* 3:71-88.

Gordon, Robert. 1986. “Folk Psychology as Simulation” *Mind & Language* 2:158-171.

Gordon, Robert. 1992. “The Simulation Theory: Objections and Misconceptions” *Mind & Language* 7:11-34.

Johnson, SC. 2000. “the recognition of mentalistic agents in infancy” *Trends in Cognitive science* 4:22-28

Lewis, David, 1970. “How to Define Theoretical Terms” *The Journal of Philosophy* 67(13): 427–446

Michael, John and Christensen, Wayne and Overgaard Søren. 2014. “Mindreading as social expertise” *Synthese* 191:817-840.

Morton, Adam. 1980. *Frames of Mind: Constraints on the Common-sense Conception of the Mental* Oxford: Oxford university press.

Nichols, Shaun and Stephen Stich. 2003. *Mindreading: An Integrated Account of Pretence, Self-Awareness, and Understanding of Other Minds* Oxford: Oxford university press.

Shanton, Karen and Goldman, Alvin. 2010. “Simulation theory” *WIREs Cognitive Science* 4:527-538

Smith, David. 2018. “Phenomenology” *The Stanford Encyclopedia of Philosophy* Stanford: Stanford University.

Zahavi, Dan. 2011. “Empathy and direct social perception: A phenomenological proposal”. *Review of Philosophy and Psychology* 2:541–558.