Erkenntnis
https://doi.org/10.1007/s10670-021-00486-7

ORIGINAL RESEARCH

Rethinking Integration of Epistemic Strategies in Social Understanding: Examining the Central Role of Mindreading in Pluralist Accounts

Julia Wolf[1] · **Sabrina Coninx**[1] · **Albert Newen**[1]

Received: 27 July 2020 / Accepted: 21 October 2021

Abstract
In recent years, theories of social understanding have moved away from arguing that just one epistemic strategy, such as theory-based inference or simulation constitutes our ability of social understanding. Empirical observations speak against any monistic view and have given rise to pluralistic accounts arguing that humans rely on a large variety of epistemic strategies in social understanding. We agree with this promising pluralist approach, but highlight two open questions: what is the residual role of mindreading, i.e. the indirect attribution of mental states to others within this framework, and how do different strategies of social understanding relate to each other? In a first step, we aim to clarify the arguments that might be considered in evaluating the role that epistemic strategies play in a pluralistic framework. On this basis, we argue that mindreading constitutes a core epistrategy in human social life that opens new central spheres of social understanding. In a second step, we provide an account of the relation between different epistemic strategies which integrates and demarks the important role of mindreading for social understanding.

1 Introduction

Humans are hyper-social beings who strongly rely on the ability to successfully interact with others. Across disciplines, researchers aim to develop an empirically adequate theory of how humans come to understand others. For a long time, theories of *mindreading* (also sometimes referred to as Theory of Mind Accounts) have

✉ Julia Wolf
julia.wolf-n8i@ruhr-uni-bochum.de

Sabrina Coninx
sabrina.coninx@ruhr-uni-bochum.de

Albert Newen
albert.newen@ruhr-uni-bochum.de

[1] Institute for Philosophy II, Ruhr University Bochum, Universitätsstraße 150, 44780 Bochum, Germany

Published online: 10 November 2021

dominated the academic discussion. According to these, social understanding relies on the *'indirect' attribution of mental states* to others, where mental states are something which cannot be directly perceived and must be inferred via theorizing (*Theory Theory*, TT) or simulation (*Simulation Theory*, ST). These accounts were criticized by defenders of *Interaction Theory* (IT), who rejected mindreading in favor of an alternative more basic, interaction-based understanding of others.

More recently, there has been a move away from such unitary theories (either TT, ST, or IT) towards an explicitly pluralist approach, according to which subjects use multiple epistemic strategies in understanding others (e.g. Andrews, 2017; Fiebich & Coltheart, 2015; Fiebich et al., 2016; Newen, 2015; Spaulding, 2018). This movement is partly motivated by ontogenetic arguments (Coninx & Newen, 2018; Newen, 2018): children acquire and unfold different epistemic abilities during development and the corresponding epistemic abilities remain in use, even after the acquisition of new epistemic strategies. Consequently, on pluralistic views, mindreading is considered to be one of several epistemic strategies used in social understanding. While pluralist theories commonly agree on this multiplicity of epistemic strategies, two central questions remain up for debate that we aim to address in more detail throughout this paper.

Firstly, what is the role of mindreading within a pluralist framework? Many pluralist accounts place significant emphasis on the alternatives to mindreading, which are thought to be sufficient for most areas of social understanding. Given that mindreading is considered to be a particularly demanding strategy, there is a tendency to view mindreading as a back-up strategy which is used only in unusual circumstances where more reliable alternatives have failed (e.g. Fiebich et al., 2016). We will call this view *unbalanced pluralism*. In this paper, we will identity three criteria that are commonly given in the literature for thinking that mindreading may be a less important component of social understanding: frequency, centrality, and reliability. On this basis, we will argue in favor of what might be labeled as *balanced pluralism*. We accept the pluralist view, but argue that (1) the empirical evidence available for considering mindreading as less frequently used is not conclusive, (2) mindreading plays an important role in everyday cases of social understanding which are central to our life and in this sense is on a par with other epistemic strategies, and (3) that it is not clear that mindreading is inherently less reliable than alternative strategies. The different epistemic strategies should be seen as complementary rather than competing approaches that are best used in conjunction.

Secondly, this leads us to the question how the different epistemic strategies are related to each other. Given that subjects use a plurality of epistemic strategies in order to successfully understand and interact with others, we need to account for the way in which these epistemic strategies are employed and combined in different settings. Pluralists might provide a compelling story concerning the ontogenetic development of different epistemic strategies, however, they usually lack a systematic explanation of which epistemic strategy will be employed by adult humans under varying conditions. Thus, what is missing is an account of how the different strategies are combined within social understanding, which demands an account of how the results of different epistemic strategies used simultaneously are weighted and how they might affect each other.

The paper proceeds as follows: in Sect. 2, we outline the main ideas of TT and ST as standard versions of unitary mindreading theories. On that basis, we briefly discuss the main objections that have motivated the adoption of a pluralistic view constituting the background for our own account. In Sect. 3, we present three possible reasons for thinking that mindreading is of lesser importance within social understanding: the frequency argument, the centrality argument, and the reliability argument. We respond to these criticisms and argue that the approach of balanced pluralism proves more adequate in the light of the significant contribution of mindreading to everyday social understanding. In Sect. 4, we outline how the multiple epistemic strategies are employed and interact across different contexts. In the absence of conclusive empirical evidence, we characterize core constraints which allows us to elaborate on the special and important role of mindreading in the interaction with other types of social understanding. Finally, we summarize our results in Sect. 5.

2 Mindreading and Pluralism

2.1 Theory-Theory and Simulation Theory

Traditional theories of mindreading assume that humans rely on the *indirect attribution of mental states* for social understanding: subjects explain or predict the behavior of others by means of indirect mindreading strategies allowing them to ascribe mental states, such as beliefs, desires, attitudes, or emotions that are supposed to underly the other person's behavior. This is alleged to be an essential part of social understanding, allowing us to obtain an understanding of others which we would not have based purely on consideration of behavior. Importantly, we cannot directly access the mental states of others in mindreading but we rely on additional cognitive processes which allow us to infer or simulate such mental states. Mindreading is most prominently tested by the false belief task (Wimmer & Perner, 1983) which children begin to pass around the age of four (Wellman et al., 2001), demarcating an important development in their mindreading ability.[1]

There are two central theories of mindreading: TT and ST. According to TT, mindreading requires inference-making based on a folk psychological theory, i.e. a system of law-like generalizations that is used to infer the mental states of others and derive their future actions. For example, Gopnik and Wellman (1992) argue that in ontogenetic development humans gradually develop a theory of mentalistic

[1] There is also some evidence that children can pass implicit false belief tasks at an earlier stage (Onishi & Balliargeon, 2005; Southgate et al., 2007). However, these tasks are currently subject to replication concerns and there is debate as to whether these should be interpreted as providing evidence of mindreading or behaviour reading (e.g. Ruffman & Perner, 2005). While we have elsewhere argued that these tasks provide evidence of an important precursor of mindreading Newen & Wolf (2020), we will not draw on this here and work with the more stringent criterion of mindreading provided by the explicit false belief task. To clarify: mindreading need not result in ascriptions, such as 'Person P believes/desires that X' but it can also be tested by checking for adequate linguistic answers which are sensitive to false or perspectival information of the other.

knowledge that relies on the same processes as scientific theories: children have expectations concerning the mental states of others, adjust the corresponding generalizations on the basis of new evidence, and thereby generate a coherent theoretical system that allows them to successfully engage in social understanding. Another version of TT by Baron-Cohen (1995) assumes that humans possess a specialized mindreading module that has been designed by natural selection for the indication of others' mental states and that comes online at a certain stage of ontogenetic development (without previously undergoing a process of expectation, falsification, and adjustment).

According to ST, we do not require a theory of interconnected mentalistic knowledge to understand others. Instead, when humans explain and predict behavior, they simulate the internal states of others on the basis of their own experience-based knowledge (Goldman, 1989, 1992). Thus, humans model the beliefs, desires, attitudes, or emotions of another person by projecting themselves into her situation while—according to Gordon's version (1986, 1992)—they adjust the projection to the particular characteristics of the concerned person. TT and ST in their early versions are monistic accounts because, according to them, a specific type of mindreading is the sole basis of social understanding. The first move away from radical unitary accounts can be found in the claim that social understanding needs a hybrid account combining, for example, the dominant simulation strategy with theory-based inferences in exceptional cases (Goldman, 2006). Nonetheless, even on the combined account, understanding others is still constrained to mindreading, i.e. the attribution of mental states, albeit carried out via two concrete strategies, and should therefore be considered a unitary account for our purposes.

Irrespective of their differences, traditional monistic theories share the view that only one main type of epistemic strategy is relevant for social understanding, namely those strategies of indirect mental state attribution that are based on cognitive processes which allow us to infer or simulate the beliefs, desires, attitudes, or emotions that underly another person's behavior. This might also include epistemic strategies developing after the age of 4, such second-order mindreading (Perner & Wimmer, 1985). In this paper, we summarize all of these related abilities under the general cluster 'mindreading'.

2.2 Direct Perception and Rule Projection

Theories of mindreading have mainly been criticized because they fail to account for less cognitively demanding epistemic strategies that subjects employ in social understanding. Such criticism has been partly motivated by ontogenetic considerations. From a relatively early stage, infants are able to socially interact with care-givers (Johnson, 2000; Reddy, 2008; Weinberg et al., 2008) in a way which suggests that they have some understanding of others' mental states. At the same time, mindreading is a fairly demanding ability which infants have not yet acquired although they already demonstrate systematic social behavior. For example, before the age of four children are able to interact with others and their environment in a coordinated and goal-directed manner and engage in joint attention (Gallagher & Hutto, 2008, 2008)

while still failing the false belief task which is used to assess mindreading. These social competences of infants in the absence of mindreading capabilities provide evidence for the existence of alternative epistemic strategies in social understanding.

Defenders of the traditional mindreading views need not be committed to denying that there are precursors of mindreading that young children make use of. These early abilities might underlie or develop into full blown mindreading abilities later on. Hence, ontogenetic arguments alone do not undermine theories of mindreading. However, the core idea of most critiques is not merely that infants employ less cognitively demanding epistemic strategies in social understanding but that these alternative strategies remain in use and remain dominant throughout life, even after children have developed the capacities required for mindreading. Researchers have highlighted that theories of mindreading systematically underestimate the role played by direct interaction in social understanding (Gallagher, 2001, 2008, 2020; Gallagher & Hutto, 2008, 2008).[2] In many situations, we are not uninvolved observers that consider the behavior of others from a third-person perspective but we dynamically interact with them (Di Paolo & De Jaegher, 2012; Schilbach et al., 2013). Particularly in such situations, we do not need to rely on inference or simulation-based mindreading because we can rely on cognitively less demanding and more direct epistemic strategies. There are two clusters of alternative strategies that are mainly highlighted by pluralists which we refer to as 'Direct Perception' and 'Rule Projection'.

First, in various social situations, we do not need to infer or simulate the mental states of others but we are able to *directly perceive* them based on social cues, such as facial expressions, eye gaze, body posture, gestures, or intonations (Gallagher, 2008, 2020; Newen et al., 2015; Zahavi, 2011). The content of perceptual experiences (Newen, 2017) or at least the content of so-called basic perceptual beliefs (Spaulding, 2015) is in itself rich enough to allow humans to access the mental states of others and to socially interact with them. For example, we can directly perceive whether a person is angry based on her facial expression without simulating how oneself would feel in the respective situation. Similarly, when jointly carrying an object, we can interact based on gestures without inferring the intentions of one another based on theory-like generalizations. Direct perception seems to unfold its relevance especially in forms of online interaction, in which we take a second-person perspective as an involved actor. Some forms of direct perception in the context of online interaction seem to be innate or very early developing (Gallagher, 2008; Reddy, 2008).

Second, in various social situations, we can directly anticipate others' behavior and react accordingly without attributing mental states at all and instead projecting behavior rules. We can predict the behavior of others solely based on our knowledge that people (both individuals or groups) follow certain social rules and show

[2] This critique was primarily driven by proponents of IT. While there are some unitary IT accounts which argue for direct perception and interaction as the only means of understanding others (e.g. Gallagher, 2001), most IT accounts have now been developed into pluralist accounts (e.g. Gallagher & Hutto, 2008).

behavioral regularities. Social interaction is often significantly shaped by norms, stereotypes, and social biases (Andrews, 2012; Spaulding, 2018) as well as by our expectations concerning the behavioral routines of an individual (Fiebich & Coltheart, 2015; Newen, 2015). For example, I might be able to predict in which manner my mother is about to react in a situation without attributing beliefs, desires, or emotions by simply projecting her former behavioral pattern into the future. Given my stereotypes concerning taxi drivers, I can predict how they are about to line up without simulating how I would behave in the respective situation. Moreover, social rules can be related to typical behavioral scripts that persons follow within certain situations (Coninx & Newen, 2018; Newen, 2015). For example, I can interact with people in a restaurant without the need of a full-blown theory of interconnected generalizations concerning how people behave in such a context. We refer to this cluster of epistemic strategies as rule projection. It involves epistemic strategies which do not require the attribution of mental states but are instead predictions of behavior based on the expectation of following social rules or the rules of an established practice. Epistemic strategies belonging to this cluster emerge in ontogenetic development when children learn to expect compliance with certain norms, natural or artificial, and protest against their violation (e.g. Rakoczy et al., 2008).[3]

At this point, one might argue that there is a third option besides direct perception and rule projection, namely the *narrative strategy,* according to which children learn to understand that people act for certain reasons through repeatedly encountering folk psychological stories. By habit and practice, these narratives enable children to internalize a loosely and non-logically connected set of shared rules and norms and to employ them in order to make sense of actions performed by others. This strategy is prominently argued for by Hutto and Gallagher (2008). It is unclear, however, how the narrative strategy is best understood. The critically relevant interpretation for us would be that narratives allow for an additional type of epistemic strategy different from mindreading.[4] Narratives would allow for predictions of the other

[3] Traditionally, mindreading and direct perception are distinguished in terms of the perspective that we adopt towards the person whose behavior we aim to explain or predict. According to this view, we act as uninvolved observers when engaging in mindreading, while direct perception comes into play in situations in which we actually engage with others (e.g. Hutto, 2008). The description of this second- and third-person nature, however, seems to primarily account for the paradigmatic situations in which we employ these epistemic strategies but are not definite of them. We may use mindreading as well as direct perception when watching another person in a train compartment and we can engage in both strategies when having a conversation with a friend. This ambivalence between third-person and second-person also persists to some extent for the alternative epistemic strategies highlighted. Rule projection may occur both within a third- person and a second-person stance. In fact, it seems that we are in everyday life often dynamically switching back and forth between a more spectatorial or engaged perspective using different kinds of epistemic strategies (see also Overgaard & Michael, 2015). We therefore do not want to focus overly on the distinction between these different perspectives. Thus, although traditional theories of mindreading themselves often emphasise the third-person nature of this approach, we do not want to commit ourselves to this. Mindreading, for our purposes, should be defined as the indirect attribution of mental states, no matter whether this attribution takes place from a spectatorial or interpersonal perspective.

[4] Alternative possible interpretations are (1) that the narrative strategy is a means of acquiring mindreading or (2) that the narrative strategy is a form of mindreading that it is based on a loosely connected set of rules. We can easily grant (1) as this would not impact our argument. Similarly, (2) would actually strengthen our argument by enriching the multiplicity of mindreading strategies.

person's behavior based on the rules, norms, and structures prescribed by the narrative for the scenario. This interpretation is supported by the claim that understanding others is not supposed to involve the designation of mental states but to rely on the exploitation of a "landscape of actions" (e.g. Gallagher & Hutto, 2008, p. 31). If narratives are understood in such a way, i.e. as providing behavioral regularities, then they would roughly do the same job as rule projection. This would also mean that our following analysis of rule projection strategies could be broadly applied to narrative strategies.

Gallagher and Hutto, however, particularly highlight that historical and situational aspects play an important role in the use of narratives, including background knowledge on an individual person or a certain group of people. However, as we will stress later on historical and situational aspects should not be viewed as alternatives to mindreading, but as components contributing to the background knowledge enriching different epistemic strategies. Thus, the main argument for a special epistemic strategy of narrative understanding can be better characterized in terms of the role of narratives for acquiring and structuring background knowledge—and not as an independent epistemic strategy—that enables different forms of social understanding and reasoning via multiple epistemic strategies. In the remaining part of the paper, we therefore think that it is adequate to primarily focus on direct perception and rule projection as the most relevant alternative epistemic strategies to mindreading.

2.3 Challenge of Pluralism

In light of these alternative strategies, various philosophers adopt a *pluralistic* view: humans employ multiple strategies to understand and successfully interact with others, including theory-based inferences, simulation, direct perception, or rule projection. Given the evolutionary relevance of social understanding and the variety of contexts in which we depend on successful social interactions, humans do not merely rely on a single epistemic strategy (Coninx & Newen, 2018; Fiebich, 2015; Fiebich & Coltheart, 2015). Hence, social understanding and interaction are more complex and diverse than unitary mindreading theories make us believe (Newen, 2015; Spaulding, 2018). At the same time, social understanding is not identical to basic interactions because there are more sophisticated epistemic strategies in use that allow us to ascribe mental states to others as described by TT and ST (Newen, 2015).

Pluralist accounts of social understanding face two decisive challenges. First, if we accept a plurality of epistemic strategies, what role does mindreading play in social understanding, particularly in comparison to other epistemic strategies? Second, in which relation do these different strategies stand to each other and when is a certain strategy activated in contrast to another, particularly with regard to adult humans to whom multiple strategies are available? We will address exactly these interconnected questions in the following two sections with the aim that they will jointly allow us to demonstrate the central role of mindreading in a pluralistic framework.

3 Mindreading as a Core Epistemic Strategy

3.1 Three Arguments of (Un)Importance

What is the role of mindreading in a pluralist framework? In general, there seems to be a focus on the importance of the epistemic strategies of direct perception and rule projection in recent pluralist accounts. Historically, emerging as an alternative to the dominant mindreading accounts of TT and ST, it seems hardly surprising that the relevance of mindreading has often been neglected, with most of the attention dedicated to carving out the role of direct perception and rule projection and highlighting their significance in our everyday lives. Mindreading, however, seems to be allocated a more peripheral role as a backup strategy for unusual situations where the alternative strategies fail. While we agree that multiple epistemic strategies are in use in social understanding, we want to push back against the idea that mindreading does not play an important role. Although not all advocates of pluralism may fully endorse *unbalanced pluralism*, there is a clear focus in the discussion on the importance of the alternatives and an explicit marginalization of mindreading by at least some authors. Our job in arguing for a *balanced pluralism* will therefore be to show where the importance of mindreading lies in a pluralist framework, and argue that this goes beyond being a mere back-up strategy for direct perception and rule projection.

Before engaging in concrete discussions concerning the (un)importance of mindreading in a pluralist account, some preliminary clarifications are needed. We need to identify and specify the possible kinds of arguments that might be put forward in order to determine which role an epistemic strategy plays in a pluralistic framework. So far, a clear systematization and discussion of such possible arguments is scarcely found in the literature. In the following, we therefore identify three types of argument that are more or less explicitly found in the literature pertaining to limit the role of mindreading within a pluralist framework:[5]

Frequency Argument: Mindreading is used less frequently in everyday life than the alternative strategies.
Centrality Argument: Mindreading is less central to human social cognition than the alternative strategies.
Reliability Argument: Mindreading is less reliable than the alternative strategies.

It should be noted that not all defenders of pluralism hold all of these criticisms of mindreading at the same time. For example, Andrews (2012) argues that mindreading is unreliable and used less frequently than alternative strategies, however, she does think that mindreading plays an important role for action explanation, even though it is not required for action prediction. Our primary goal is not to prove that

[5] This distinction of three possible arguments for or against the central role of mindreading has been introduced in Wolf & Coninx (2021) with the discussion of the corresponding arguments differing and complementing each other at key points.

a particular author holds a particular position, but to show a general tendency in the debate regarding the role of mindreading in pluralistic theories and to identify relevant arguments for this. This can hopefully provide the basis for further discussion regarding the role of different epistemic strategies, regardless of the position that an individual author might hold.

3.2 Frequency Argument

The *frequency argument* states that mindreading is used less frequently than alternative strategies in everyday interaction with other people. For example, as Gallagher outlines at length:

> Of course, this is not to say that we never use conscious simulations or theoretical inferences. It may be the case that confronted with some strange or unaccountable behavior I do try to understand the other person by running a simulation routine or by consciously making a folk-psychological inference. This type of process may also occur in specialized situations; for example, when we are playing poker, or practicing psychotherapy, or, as Overgaard and Michael (2015) suggest, on a first date. Just such instances are telling, however, since they are in fact special circumstances or specialized cases. Moreover, such cases tend to stand out in their rarity. (Gallagher, 2020, p. 81)

To provide an ultimately convincing argument that relies on the frequency of use one would need to define criteria of what exactly counts as "always used", "mostly used", "sometimes used", or "rarely used", for example.[6] Based on this conceptual clarification, determining the frequency of an employed epistemic strategy would then be an empirical matter. However, it seems very difficult to clearly individuate different episodes of social understanding in a rigid manner, count them across various real-life conditions, and evaluate them in relation to a pre-fixed benchmark of different frequencies of use.

Given these problems, most authors seem to rely on experiential evidence of everyday engagements in social understanding where we usually do not phenomenally experience ourselves as engaging in mindreading (e.g. Gallagher, 2020, pp. 80–82). This argument seems problematic in the absence of further empirical evidence, however, as the experienced frequency of mindreading may vary between individuals. Furthermore, it depends on the accuracy of introspection in detecting mindreading which may not be given. Similar problems are found with regards to Fiebich et al.'s (2016) argument that mindreading is less frequently used because it is more demanding than alternatives and not required for the majority of our social interactions. However, as we will show in the next section, the alternative strategies are perhaps more limited than some think. While we do not aim to argue that mindreading is on the whole the most frequent, we do want to point out that the extent to which one encounters situations requiring mindreading is also subject to individual

[6] Thanks to one of the anonymous reviewers for stressing this point.

variation and therefore empirical research would be necessary to determine how frequent these situations actually are. This means that both the arguments provided are to some extent dependent on empirical verification.[7]

Currently, there is very little evidence on this issue. One exception are the studies of Malle et al. (2005, 2007; Korman & Malle, 2016) which distinguish and compare three types of explanation of behavior: (1) reason explanations, i.e. ascribing beliefs or desires as core causes of the agent's action which can be classed as mindreading explanations, (2) causal history of reason explanations, i.e. referring to causal aspects external to the agent's beliefs or desires (e.g. an agent's disposition, personality, upbringing, culture, or context); and (3) enabling factor explanations, i.e. referring to those aspects that explain how it was possible that an agent's intention turned into a successful action. The core result of the comparison is that reason explanations are the most frequently used type of the three across several studies cited above. This interpretation is defended in more detail by Peters (2021).

Taken together, there are remaining conceptual issues on how to define the relevant frequency of use of different epistemic strategies in a pluralist framework and empirical issues on how to exactly measure such frequency. The current evidence base concerning the frequency argument against mindreading is far from conclusive and the rare empirical evidence may actually indicate the opposite. The available quantitative studies support our claim about a special role of mindreading but we grant that more empirical research is needed to clarify this feature. Even if we were to accept, for the sake of argument, that mindreading may in principle occur less frequently than the alternative strategies, this alone is no convincing argument to consider mindreading less important than alternatives. For the remaining part, we will therefore address the other two arguments.

3.3 Centrality Argument

Our main target will be to argue for the importance of mindreading in reply to the *centrality argument*. That is, we argue that mindreading is a core epistemic strategy within social understanding because various spheres of our social life are only accessible via mindreading. While these situations may not be the most frequent, they are particularly central to our existence as human beings. Unlike the frequency, determining the centrality of an epistemic strategy is not a quantitative matter: it is a question of which situations of human life matter most to us. At least in some pluralist accounts we find authors drawing on intuitively central cases of social understanding and arguing that mindreading is not required for these:

[7] As an additional strand in the frequency argument, one might argue that mindreading is only employed when we take a spectatorial stance while we mostly engage with others from a second-person perspective (e.g. Hutto, 2008, p. 13). Although mindreading is not only employed from a third-person perspective, it surely is when we do so (for more detail see footnote 3). This alone is, however, not sufficient to determine the frequency of mindreading. For example, Overgaard and Michael (2015) have indicated that the spectatorial perspective is not unnatural for us: otherwise it would be hard to explain why we watch reality TV or talk about others, preferably when they are absent.

> In general, pluralist approaches contend that third-person mental state attributions that may involve theory or simulation do not lie at the heart of everyday social understanding and often come into play peripherally in ambiguous or unfamiliar contexts that make individuals puzzles about the other person's behavior. (…) Other factors such as trading second-person narratives; being sensitive to environmental contexts, norms, habits, social conventions; and having knowledge of character traits of familiar individuals are far more central. (Fiebich et al., 2016, p210)

Our approach to this will be twofold: firstly, we will present a number of prototypical situations of social understanding where we argue that mindreading is necessary. We will primarily refer to examples that various authors, including those that seem to oppose our view, accept as cases of mindreading. Thus, our aim is not primarily to show that these are cases of mindreading, but that they are far more central to our social life than assumed. Secondly, we will consider the case of people with disorders impairing their mindreading abilities and provide evidence that this impacts their social understanding. While evidence from people with disorders is often difficult to interpret conclusively, taken together with the previously discussed examples we think that they provide an indication in a similar direction and at least show a potential area where the centrality of mindreading could be further assessed empirically.

3.3.1 Three Central Cases of Social Understanding

Our argument for the centrality of mindreading requires two steps. In a first step, we must show that the alternative strategies to mindreading are limited and that therefore there are situations in which mindreading is required. The second step then is to show that these situations are in fact central to human life. Evaluating the centrality of an epistemic strategy is not merely an empirical matter but depends on what we consider important for us as human beings. With the following examples we hope to illustrate that our social life would be substantially different with respect to some of its core aspects if we could only rely on direct perception and rule projection.

In the previous section, we saw that pluralists offer two main clusters of alternatives to mindreading: more direct means of perceiving the other's mental states which do not require inference or simulation and forms of rule projection which allow us to predict another person's behavior without attributing mental states to them at all. However, the corresponding epistemic strategies, taken on their own, are quite limited.

It has long been considered uncontroversial that we are able to directly perceive mental states like basic emotions such as anger and fear or basic intentions such as changing direction. If we have sufficient background information, for example, if we know the context and the person in question, we might be able to engage in smart direct perception allowing for access to more complex feelings and intentions (Gallagher, 2008). Still, these direct forms of social understanding provide only limited access to more complicated thoughts and affective entanglement and are necessarily limited to the respective scenario, leaving many crucial situations of social

understanding uncovered. The same is true of rule projection strategies which do not involve the attribution of mental states but only behavior prediction in line with certain rules applicable in the relevant situation. For example, with the rule projection strategy, I can only predict that a soccer player who has a bad fall will scream out loudly and hold their knee, but not what their mental state will be. Rejecting mindreading in favor of rule projection strategies leaves us either with a rather radical, almost behaviorist stance or, in combination with direct perception, it allows at best for registering fairly simple mental states according to rules in the concrete situation but excludes flexibility in our interpretation of the behavior of those who we aim to understand.

It might be argued that this is not a problem for defenders of unbalanced pluralism, as they do allow that mindreading plays a role in extraordinary circumstances. However, as we shall see, such circumstances often characterize central situations of social understanding. In what follows we shall therefore discuss some key examples of situations in which mindreading is required and argue that—while they may not be the most frequent—they are nonetheless central situations of human social understanding. We will concentrate on three types of cases that are commonly accepted as requiring mindreading (see prominently Gallagher, 2001, p. 92): (1) cases in which the behavior of a person is deviating from the expected, puzzling, and/or suspicious, (2) cases in which we engage with someone we are not (yet) familiar with, and (3) cases in which we talk or think about an absent person.

First, epistemic strategies of direct perception and rule projection fail as soon as a social situation becomes more complex and variable, but this complexity and variation is almost unavoidable in certain cases of social interaction (see also Wolf & Coninx, 2021). Imagine a couple having dinner after work. Direct perception might be sufficient for the husband to recognize that his partner has something on his mind, but mindreading is required to consider the exact reasons for the partner's behavior. Direct perception tells us that something is wrong but we need mindreading to determine what is wrong. Rule projection cannot replace this, since this is about understanding the mental states of a specific individual in a concrete life situation whose behavior might deviate from the known general rules of behavior, even from the typical behavior of the respective person themselves. That is, mindreading comes into play here as rule projection has failed and direct perception comes to its limits.

It is readily allowed by some defenders of pluralist views that such situations in which we are presented with unexpected or puzzling behavior of others do in fact require mindreading (Gallagher, 2020). What might be more controversial is the question of whether such situations are central to social understanding. For example, romantic couples may strive to adjust their behavior to make themselves better mutually understandable which would in principle reduce the necessity for mindreading.[8] In the example given above the problem may not lie in lacking comprehensibility, but rather that more detail is required than the alternative strategies allow for. Direct perception and rule projection might only function as indicators that

[8] Thanks to one of the anonymous reviewers for stressing this point.

further investigation is needed. Still, it might seem that such situations do not necessarily demand for mindreading as the husband in the previous scenario could simply ask his partner what is wrong. What this ignores, however, is that there may be situations in which the other person does not want to be comprehensible, for example, because they are embarrassed and do not want to share their problems. A more extreme example of this would be cases of outright deception, but even where no malice is concerned we may intentionally or unintentionally fail to increase comprehensibility. In these situations, direct questions are also unlikely to bring clarity.

What makes mindreading central for these situations is not that we are always distrusting of the other or trying to go beyond what the other person is directly making available to us, but rather that we have the *potential* to go beyond this. Mindreading allows us to have an understanding of others that goes beyond their straightforward behavior or directly perceivable mental states and it means that we are not naïve victims of social interaction.[9] Even if humans mostly act cooperatively in everyday life, the cases in which we don't are often of utter relevance. These are not just cases of poker games in which we seek to understand the mental states that others aim to hide. The potential to go beyond the immediately given might come into play in critical moments of a relationship (e.g. aiming to care for another person when they are unable or unwilling to explicitly share their feelings) or areas where it can be dangerous for our physical well-being or socio-economic status if we are not able to protect ourselves from deception (e.g. a colleague who tricks us to get promoted faster, a partner who cheats on us, or a person who tries to cause us financial or physical harm).

Second, mindreading is often argued to be used in novel situations, as the alternative strategies come to a limit when the other person is very different to us or behaves in an unfamiliar manner (see also Fiebich et al., 2016; Gallagher, 2020; Hutto, 2008). A common example given is that of people who come from different cultures or suffer from disorders impairing abilities of social interaction. Here it might rightly be objected that such scenarios are quite unusual, although becoming more frequent in our increasingly globalized world and integrative society.[10] However, we also need to employ mindreading when getting to know a new person, especially in contexts where we need to evaluate whether we would like to maintain a long-term relationship. This might include situations such as a date, a social event at a friend's place, or meeting a new work colleague. In the situations just mentioned, we usually do not only aim for smooth interaction in the very situation but we also want to understand and predict the behavior of the person in the future. Therefore, we are interested in getting to know the world view, life goals, or emotional depth of a person which requires the ascription of mental states, such as beliefs, desires, or non-basic emotions. To reach this aim, we need mindreading: With direct perception we can only register basic emotions and goal-directed actions in the very

[9] This might even apply if we are not good at detecting deception, as our point is that mindreading is the only strategy to successfully prevent deception, provided that we detect it in the first place.

[10] These types of situations also extend to children, whose behavior differs from ours (for more detail see Wolf & Coninx, 2021).

situation. Furthermore, with rule projection in combination with knowledge about certain social role we might be able to develop some expectations about convention-based behavior which improves our social interaction. These expectations, however, remain still underdetermined for many situations in which we aim to gain knowledge about a particular person, including their unique beliefs, desires, non-basic emotions, and so forth. If we can use mindreading, we actually do so because the benefit of improving predictions of behavior, especially concerning individual differences and predictions going beyond the immediate situation is so significant.

Overgaard and Michael (2015) have recognized two important features of such situations which highlight the relevance of mindreading. First, it is often inappropriate to constantly ask about how a person is feeling or what they are thinking in such situations. Second, these kinds of situations in which we get to know a new person are not cases in which the other strategies have failed: mindreading is not the back-up for direct perception or rule projection but the default strategy to set up relationships. Only mindreading enables us to understand others with respect to their future related plans, intentions, and goals. When getting to know someone, especially with regard to those people who are important to us, we are particularly interested in the other person as an individual going beyond the here and now. This individualized approach which allows us to learn substantial new things about another person including their mental states is something which we require mindreading for. Interestingly it is granted by some critics of the centrality of mindreading that these situations require mindreading (Gallagher, 2001, 2020), the question is whether these cases are central to our social lives, independent of their exact frequency. In our view, getting to know new people and deciding who we want to engage with in long-term relationships are critical scenarios of our social lives in personal and professional contexts which may become central turning points in our lives.

Third, we necessarily engage in mindreading when talking about an absent third person; for example, when deliberating with our siblings about the birthday present for our parents. It seems that if we need to take a third-person perspective in dealing with the mental states of others, mindreading is the only strategy available to us. While we might accept that these are cases of mindreading, one might still doubt whether these cases are more than peripheral to our social lives (e.g. Hutto, 2008, p. 13). However, it seems hard to imagine that we do not speak about other people in the third person at key moments and attribute certain mental states to them. For example, we might talk to our siblings about what kind of gift would make our parents the happiest, we might talk to other employees about which applicant would be the best fit for the company, or we might talk to a friend about what another friend must have thought and felt at the party yesterday in order to explain their unusual behavior.

Further, it seems important to note that we do not only engage in mindreading from a third-person perspective when talking about someone, but also when we think about what someone else would think or feel if we behaved in a certain way. Especially where people who are important to us are concerned, our engagement with them and their interests does not only take place when we are interacting with them directly, but also when they are absent. It seems to be an important part of generating and maintaining relationships that we think about how our behavior affects

others and we adjust it accordingly. This might range from cases as mundane as wondering whether a friend will be happy about a gift, to thinking about how a partner will feel about a particular life choice we make.

Finally, it should be noted that in these cases the other epistemic strategies do not seem to break down, but mindreading allows us to tap into a different social sphere in which we can consider the mental states of others in their absence. Direct perception does not seem to help and rule projection is also of limited use as it is precisely the specific mental states of a person that are being considered rather than merely their obvious behavior. Again, one might object that the person in question could of course be asked about their mental states. However, we may not want to do this in many cases, either because we want to surprise the concerned person, or because we want to decide first whether to tell a person about an issue, or because it would be against social conventions.

3.3.2 Down Syndrome

As a second line of argument, we aim to outline the importance of mindreading in reference to subjects whose abilities in this area are impaired. If mindreading is used solely in rare exceptional cases, restrictions in mindreading abilities should have no significant impact on the social understanding and interaction of concerned subjects. As we shall see, there is quite some evidence that individuals with impaired mindreading abilities are also limited in their social understanding. Classically, autism is used as an example of people who lack mindreading abilities and therefore display significant social impairment. However, the evidence in this regard is less clear, especially as people with high functioning autism seem capable of some kind of mindreading (Luckett et al., 2002; Schuwerk et al., 2015). It is interesting, however, to contrast this with cases of Down Syndrome, which have received less attention in philosophical debate.

Like other people with intellectual disabilities, people with Down Syndrome show significant deficits in complex social tasks related to mindreading (e.g. Thirion-Marissiaux & Nader-Grosbois, 2008). This is particularly salient in tasks requiring the adoption of different points of view on the same situation, such as in visual perspective taking. At the same time, people with Down Syndrome are often perceived as very sociable. At first glance, this case study therefore appears to confirm the minor role of mindreading posited by defenders of unbalanced pluralism: people with Down Syndrome often have an outgoing personality, show basic emotional sensitivity and empathy, and in direct interaction with others they typically perform emotionally expressive behavior, providing the basis for successful interactions in multiple everyday situations (Fidler et al., 2008).

People with Down Syndrome are normally not able to develop a situationally independent social understanding accounting for others' deviating intentions or long-term plans as this understanding requires perspective taking, mental time travel, and executive planning. At the same time, we see that people with Down Syndrome show limitations in the generation and maintenance of social interactions. First, people with Down Syndrome are restricted in their ability to attribute thoughts, preferences, and intentions different from their own, especially when they are unfamiliar

with the person in question (Giaouri et al., 2010). This carries the risk of inappropriate interactions with others, such as approaching strangers in an indiscriminate manner (Porter et al., 2007). Second, it seems more difficult for people with Down Syndrome to lie to or manipulate others, as can be seen in their poorer performance in deception taks (Amadó et al., 2016). Third, against common assumptions, people with Down Syndrome often struggle to develop sustained friendships with others (Cebula et al., 2010; Wishart, 2007).

Most convincing evidence for our purpose would be provided by studies that control for the deficits of people with Down Syndrome with regard to different epistemic strategies. Unfortunately, such studies do not exist, at least to our knowledge. Moreover, we are aware that a strict distinction between impairments with respect to different epistemic strategies of social understanding is hardly possible. These often go hand in hand and are conditioned and amplified by more general cognitive and motoric impairments (e.g. Cebula et al., 2010). For example, there are deficits in more basic social competences. People with Down Syndrome show difficulties in recognizing and adapting their behavior in response to facial expressions, especially associated with fear and anger (Cebula et al, 2010; Goldman et al., 2018; Wishart, 2007). An inappropriate approach to strangers may be fostered by difficulties in direct emotion recognition as well as mindreading (Porter et al., 2007). Thus, it would be inadequate to characterize basic social understanding as protected from deficits altogether, although these abilities are sufficient insofar as they enable an adequate short-term, situation bound social interaction.

Our aim is not to show that the impairments in the case of Down Syndrome provide the ultimate evidence for the social cognition of neuro-typical people. Moreover, we are aware that the current empirical basis is not fully conclusive. However, we think that the available data provide an indication for a potentially promising area of research where the centrality of mindreading can be assessed empirically. The case of Down Syndrome indicates that the inability to attribute mental states to others, which seems to be most affected, is at least related to the inaccessibility of central spheres of social understanding. In particular, this concerns the initiating and maintenance of long-term relationships which seem to require a deeper understanding of the other person. Further, this affects the recognition of intentions that others may try to hide from us or the manipulation and deception of others—aspects that are often neglected and that we have highlighted as central in the previous section.

Taken together, the examples in Sects. 3.2.1 and 3.2.2 show that mindreading does not only occur in the periphery of our everyday life social understanding. We do not assume that mindreading is the *only* epistemic strategy used in many prototypical situations without also making use of forms of direct perception or rule projection. However, mindreading is central to our social lives. It comes into play in critical situations when direct perception and rule projection have failed (e.g. deception) and it opens up new social spheres that would otherwise not be accessible (e.g. when getting to know new people or deliberating about their mental states when not directly engaging with them). Although they may be less frequent, these situations prove central to the generation and maintenance of personally and professionally significant relationships.

3.4 Reliability

The *reliability argument* is the last argument which is often raised against mindreading, claiming that it is unreliable and therefore cannot be the strategy which usually underlies our social understanding:

> A second reason to reject the idea that we typically attribute beliefs to others in order to predict their behavior is that attributing beliefs isn't a very accurate way of predicting behavior—and we are pretty good at predicting quotidian human behavior. (Andrews, 2017, 119)
>
> That is, we may be forced to make third-party predictions and explanations of actions precisely in the sort of cases in which we do not know what to expect from others or when we cannot engage with them directly. But, for this very reason, these sorts of approaches are bound to be, on the whole, much less reliable than our second-person modes of interaction. (Hutto, 2008, p13)

When it comes to reliability, as with frequency, we face conceptual as well as empirical issues (see also Wolf & Coninx, 2021). First, it may seem hard to precisely define when something counts as reliable in terms of an exact number of how often an epistemic strategy needs to be successful rather than unsuccessful. Second, as Westra (2020) has argued, we do not currently have any good means of assessing the reliability of mindreading: the only way in which we can assess the reliability of mindreading (especially for more complex mental states) is by determining whether it fits with self-report. There are serious concerns about the reliability of self-reports of mental states, however, which make this assessment problematic. Does this mean that nothing can be said about the reliability of epistemic strategies at all? Not necessarily. Andrews (2017) and Hutto (2008) provide reliability arguments that are based on more general theoretical principles and thus avoid the previous conceptual and empirical challenges. It is these criticisms we focus on in the following discussion.

First, Andrews (2017, p. 119) argues that there is no one-to-one relation between mental state and behavior. This means that a multitude of predictions would be compatible with a given behavior and in light of these many possibilities the accuracy of mindreading is likely to be low. The alternative strategies of direct perception and rule projection are, however, equally subject to these limitations, meaning that mindreading is no worse off than the alternative strategies. There is no one-to-one correspondence between bodily cues and a certain mental state to be directly perceived or between a certain interactive situation and a certain rule or norm to be employed.

Direct perception is limited to superficial and basic properties and the same facial expression, gesture, body posture, or intonation might be associated with multiple mental states. For example, seeing someone smiling in many cases does not provide a reliable indicator for joy given that smiling is considered as an expression of politeness in many cultures or used in sales contexts to communicate approachability. Similarly, rule projection strategies struggle with individual and contextual differences. Norms concerning social groups or situations might lead to errors due to the fact that people do not always adhere closely to the correspondingly predicted behavioral patterns and behavioral patterns may vary substantially between different people. That is, the same situation might require interaction in relation to multiple

norms or scripts. Engaging with a certain person, we might rely on rule projection concerning their profession, concerning the situational script, concerning the social norms we have developed for such individuals, and so forth.

Second, Hutto (2008, p. 13) argues that mindreading is less reliable because it is employed in those cases in which we have no information from direct perception and have little background on what to expect. While we agree that we must rely on mindreading in such situation and that the corresponding indirect ascription of mental states might be quite error-prone, this does not tell us anything about the reliability of mindreading as such. The reliability of all epistemic strategies equally depends on various factors, including the reliability of the background knowledge. For example, direct perception often depends on our knowledge of important background information and it is only as reliable as this background information. For example, if we see someone crying while having missed the preceding marriage proposal, we might erroneously think that the person is sad even though the reverse is true. One could argue that direct perception and rule projection prove to be much more reliable if we have access to more information. This is true, but the same might be in principle true of mindreading. There are less misunderstandings based on mindreading if we know the person we aim to understand well or have information about the situation they are in.

Taken together, the first impression that mindreading is less reliable than other epistemic strategies may be simply that we use this strategy especially in conditions where we can rely on little background knowledge, for example, when we meet a new person and need to gain information on such person in the first place. However, this does not mean that the low reliability is due to mindreading itself. Moreover, as we have outlined before, there are different cases in which mindreading is employed and, so far, we lack a convincing reason that mindreading in itself is unreliable in all these different cases. Thus, the criticism by Hutto seems to rely on a too narrow understanding of when mindreading comes into play.

Finally, we want to highlight that all epistemic strategies are likely to have their limitations given the multiplicity of social situations they might be applied to. All of them can suffer from relatively low reliability when taken on their own. One central idea of some pluralist approaches is not only that there are different epistemic strategies in social understanding but that they are complementary and reciprocally interacting strategies that are often simultaneously used (Spaulding, 2018; Westra, 2019a). This interplay between the different epistemic strategies may allow for an increased overall reliability but, more importantly, what this indicates is that mindreading should not be seen primarily as an alternative to other epistemic strategies, but as a component used in conjunction with them. Of course, there may be some scenarios where we really are just interested in online interaction or behavioral prediction or mindreading. However, it might seem plausible that social cognition is most *effective* when the different epistemic strategies are combined and we flexibly switch between them depending on the requirements of the situation.

Having thus argued that mindreading has an important role in a pluralist framework, the remaining part of the paper will strive to provide a more structured approach to the question when the respective epistemic strategies are employed and how they interact with each other. This allows us to further elaborate on the

important role of mindreading in the interaction with other types of social understanding and provides support for the previously outlined assumption that the different epistemic strategies are not to be primarily studies in isolation but in their integration.

4 The Role of Mindreading in Balanced Pluralism

In the following, we aim to present a theoretical framework that allows us not only to grant a plurality of epistemic strategies, but also to answer pressing questions about when these strategies are activated and how they interact to enable a flexible understanding of others. In order to develop our idea, we first argue that social understanding requires an integration of epistemic strategies and background information. This can be systematized along three central and idealized types of social understanding resulting from the combination of the previously introduced clusters of epistemic strategies with the kinds of background knowledge typically employed (Sect. 4.1). Subsequently, we describe the dynamical role of these types of social understanding in the context of bottom-up and top-down activation (Sect. 4.2). Thereby, we aim to account for the flexible change between and combination of the different types of epistemic strategies and background knowledge enabling humans to reliably understand and interact with others, once again highlighting the important role that mindreading plays in social understanding.

4.1 Idealized Types of Social Understanding

Social understanding involves the employment of at least one epistemic strategy from the three clusters previously introduced: direct perception, rule projection, or mindreading. However, in order to characterize different types of social understanding, it appears important to not only refer to these clusters of epistemic strategies but to also consider different types of central *background knowledge* that feed into our understanding of others. Independent of the epistemic strategy employed, we rely intensely on prior knowledge about situational requirements, social roles, and individual subjects in order to develop and execute the relevant strategy.[11] In the literature we find three idealized types of structured background knowledge, namely (1) *situation models*, (2) *groups models*, and (3) *person models*.[12]

[11] This idea has been defended in more detail in Newen (2018) and Coninx & Newen (2019). Therein, two basic background models are distinguished: situated models and person models, where person models can apply both to social status groups or individual person. In this paper, these two kinds of person models will be kept separately labeled as group or person models where the latter focuses on individuals only.behavioral regularities of an individual

[12] The models of situations, groups, and persons also unfold ontogenetically and they are systematically enriched by new information received. Thus, these models are not rigid but dynamically developing, fostering flexible and reliable social understanding and interaction based on accumulated social experiences (Coninx & Newen, 2019).

Situation models integrate information about prototypical situations and include expectations about patterns of behavior regularly performed in these situations. In combination with direct perception, we receive a *perception-situation understanding* of others which is already implemented by taking and handing over an object in a minimally cooperative context. We understand the other as taking over a basic situational role, e.g. being the one who hands over an object, and interact with them by directly perceiving each other's behavioral intentions. If we combine situation models with the epistemic strategy of rule-projection, then we receive a *rule-situation understanding* which is typically realized as forms of interaction along behavioral scripts: we engage with others by predicting their behavior in line with the relevant rules (e.g. rules for visiting a restaurant). The third option is to engage in *mindreading-situation understanding*. For example, in the example of the injured soccer player (see Sect. 3), we need mindreading along with background knowledge concerning their delicate situation to understand that the person is not only in pain but also in a state of intense fear.

We have described the three idealized cases of combining *situation models* with three types of epistemic strategies. The same flexible combination is possible with respect to the other kinds of background knowledge. *Group models* integrate information that is typically associated with a certain group of people instantiating a specific social role (e.g. taxi drivers, school teachers, lawyers). For example, this includes background expectations concerning the behavior that persons of the relevant group are supposed to show (e.g. how school teachers interact with pupils and with parents). *Person models* integrate characteristics of specific individuals, such as their regular behavior (e.g. to do sports every Monday) as well as their mental states (e.g. specific preferences, emotions, moods, beliefs, desires).[13] Social understanding typically relies on the employment of an epistemic strategy in conjunction with prior knowledge about relevant situations, groups, and individual subjects. This gives rise to a variety of *types of social understanding* resulting from the combination of certain types of epistemic strategies with certain types of background knowledge, as shown in Table 1.

To simplify the presentation and discussion of the plurality of types of social understanding, we highlight three idealized types of social understanding which result from the combination of one cluster of epistemic strategies with one dominant type of background knowledge (see also *[authors]*).

Perception-Situation Understanding: This idealized type of social understanding is realized by direct perception. Epistemic strategies in this cluster often correlate with the activation of simple situation models which involve *short behavioral programs* triggering adequate complementary action or basic coordinated cooperation within a particular situation based on the perception of basic mental states or motor

[13] For our purposes we can leave it open how exactly the relation between behavioral regularities of an individual and their mental states should be spelled out. Furthermore, person models and group models typically have overlapping contents because much of what we know about a person are social group roles, e.g. we know that a person is a medical doctor and a parent of two children. Thus, we expect typical features of their mindset, namely that they like to heal people or that they love their children. Furthermore, we may know individual mindset features, e.g. that they like modern art.

Table 1 Combinations of types of epistemic strategies and types of background knowledge

	Situation model	Group models	Person models
Direct perception	Perception-situation understanding	Perception-group understanding	Perception-person understanding
Rule projection	Rule-situation understanding	Rule-group understanding	Rule-person understanding
Mindreading	Mindreading-situation understanding	Mindreading-group understanding	Mindreading-person understanding

intentions (e.g. handing over an object, jointly carrying an object, or shaking hand). These simple situation models might be available from early on in childhood. However, this type of social understanding can only be effective in simple and highly standardized contexts in which deviations and individual differences are not decisive for the success of interaction.

Rule-Group Understanding: This second idealized type of social understanding concerns the expectation that others follow rules which are constitutive for members of a certain group according to a group model (Newen, 2015). For example, when we are introduced to another person as a school teacher, we expect her to act according to her social role as a school teacher in the classroom (e.g. teaching, giving homework, running exams) and, in addition, we assume that she acts in a particular manner in other situations as well (e.g. talking about school or educating people about certain topics). Hence, we activate our stereotypes of school teachers and expect the relevant person to act according to the behavioral profile which constitutes these stereotypes and interpret their behavior according to these rules.

Person-Mindreading Understanding: The person model of an individual subject entails a large unity of integrated information about such an individual subject including physical properties like body height, weight, perceptual information concerning, for example, the face, as well as information concerning stable social relations or the cognitive profile of the person. When we are interested in understanding a particular human as an individual person, there is a contextually relevant part of the extended person model which is especially important, namely information concerning the life plans, hopes, desire, emotions, and beliefs which we may call the *mindset* of the relevant person in a certain social embedding relative to a situation. This activation enables us to attribute an unconventional set of beliefs, desires, and other mental states, such that we can explain and predict the behavior of the person, even if this behavior deviates from our expectations related to a social role, for example. The combination of mindreading and mindset is important not only because background knowledge of mindset is often used in conjunction with mindreading, but also because it is mindreading that initially allows us to determine the individual mindset of a person. This is the case for colleagues at work, friends, or family members. In all these cases, we need mindreading to create and update a telling mindset of a person that we aim to understand and interact with over a longer time period in order to lead healthy relationships with a partner or to cooperate in a professional context, particularly in our digitalized society (see Sect. 3).

It is important to highlight that these are three idealized types of social understanding which are often used in parallel and with some intertwinement. We explicitly allow and expect that several types of social understanding are used at once. We use all information available in a situation in order to come up with the best possible social understanding. For example, in one and the same dinner situation, we can understand the behavior of the people involved making use of situation-perception understanding which allows us to perceive that one person is angry and not talking to others as well as of group-rule understanding which enables us to notice that after an opening toast all guests have to shortly introduce themselves to the table. On this basis, we might start to also activate the relevant person models for groups. Furthermore, we rely on person-mindreading understanding if we get involved in a

conversation with our direct neighbor at the table: we activate a person model for the very individual including information about the person's plans for upcoming summer and the like and ascribe certain mental states based, for example, on theorizing what a person with such plans might think or hope for. This illustrates that in many situations we can rely on all or at least more than one of these types of social understanding at once. These idealized types of social understanding remain of course examples of the complexity of possible combinations of clusters of epistemic strategies and kinds of background knowledge (see Table 1).

4.2 The Dynamics of Social Understanding

Let us summarize the plurality of social understanding: a type of social understanding is typically based on the activation of an epistemic strategy together with a model of background knowledge related to a situation, group, or individual. The resulting types of social understanding can be activated separately or in combination with other types of social understanding in one context. It is important for humans that we have a plurality of types of social understanding available because they enable us to make flexible use of them given different contexts and social challenges. According to our knowledge, there is little research on when the respective types of social understanding are employed or which types are combined in a situation, which remains an open challenge for empirical research.

Can we nevertheless make progress in constraining the dynamics of using a certain type of social understanding? We think that this is possible relying on a Bayesian framework. There is increasing evidence that social understanding can be adequately described as a kind of *Bayesian weighting process* (see also Bruin & Strijbos, 2015; Westra, 2019b). This involves processes which automatically evaluate the accessible types of social understanding and select and activate those which produce the most adequate social understanding and interaction given the challenges in the situation (review by Brown & Brüne, 2012).[14] Bayesian statistical inference is a method for generating posterior expectations for potential outcomes/hypotheses based on new evidence/information as well as prior expectations (potentially shaped by even earlier experiences). More specifically, this statistical method can be used to determine the probability of a certain hypothesis, given predetermined assumptions (i.e., the prior probability of that hypothesis as well as the likelihood of the actual outcome in the light of that hypothesis). Based on such Bayesian approach, different

[14] It is now fairly well established that e.g. our sense of agency is grounded in a Bayesian weighting process (Synofzik et al., 2008a, 2008b). This is a form of self-understanding and it is plausible that understanding others is closely connected to self-understanding (Decety & Sommerville, 2003). Therefore, understanding others is probably also a Bayesian process. We propose that the implementation of the Bayesian process can best be described within the framework of predictive processing (Friston & Frith, 2015) but we remain open to re-describe the basic weighting and evaluation process in more general Bayesian terms.

types of social understanding can be activated in two ways, namely either bottom-up or top-down.[15]

Bottom-up activation of social understanding is input-driven: in some standard situation of social interaction, we process central social cues, such that we see another person approaching with a smile and outstretched hand and correspondingly expect a handshake. Bottom-up activation can be based on a single social cue, e.g. facial expression or biological motion, but it typically relies on the integration of a multiplicity of social cues at once. In the case described, the perceived cues trigger the basic perception-situation understanding with a behavior program of greeting each other. In bottom-up activation the reliability of social cues is rated rather highly and perception-situation understanding might be sufficient in some contexts. However, we can easily imagine a situation in which the corresponding expectations are not fulfilled and need to be updated based on the employment of a more demanding type of social understanding, such as rule-group understanding. For example, imagine a Western person approaching a Japanese business partner. She expects a handshake, but the business partner does not join in due to different conventional rules and just bows. She might instantaneously be reminded about the intercultural differences and not interpret this behavior as unfriendly. A slightly modified scenario gives rise to the level of mindreading-person understanding, e.g. if someone notices that the person approaching with a smile and an outstretched hand is a personal enemy with bad intentions. By activating the person model of this individual, one overwrites the behavioral disposition to greet by handshake and may decide to treat him only minimally polite with a formal 'hello' ignoring the greeting gesture.[16]

Although the bottom-up perspective primarily starts with the perception-situation understanding and may not always need an activation of a higher level of social understanding, we nevertheless rely on higher levels of social understanding including mindreading-person understanding, especially in those situations in which the other epistemic strategies break down. That is, for example, when the behavior of others is unexpected, deviates from the norm, or we suspect deception (see also Sect. 3.3.1). We usually aim for a multi-level social understanding of the same person in the situation.

Top-down activation of social understanding is expectation driven: Top-down activations are driven e.g., by expectations about the mindset of a person, i.e. the set of beliefs, desires, hopes, fears, intentions, and the like that characterize such a person. It is a typical feature that the reliability of the information in the background model (in this case, the person model) has high values and is not easily modified. This results in an interpretation of the social cues and even behavioral scripts in the light of the mindreading-person understanding. For example, if we know that a closely related person is going through a difficult episode of their life and already

[15] We do not understand this in a neural processing way but according to psychological levels distinguishing an input level and one or several levels of hypothesis or expectation. Then we highlight the difference in the initiation and dominance of the information for the processing of social understanding.

[16] Which epistemic strategy we use from the cluster "mindreading", i.e. inference or simulation-based strategies, can depend on different factors. For example, in interactions with people who are very similar to us, we may rely more on simulation and for people dissimilar to us on inference.

expect certain mental states to be present (e.g. sadness), we actively search for cues indicating emotional difficulties in order to adequately support her. This might include deviations from the behavior typical for the person as an individual or member of a social group as well as changes in facial expression or movement. In the case described, tears are not read as possible expressions of positive or negative feelings, but rather, driven by the mindreading-person understanding, are seen as confirmation of our expectations that the person is emotionally vulnerable.[17]

At least under two standard conditions the described form of top-down activation is employed while our expectations concerning the mindset of the person shapes the employment of rule-group and perception-situation understanding, given that these are simultaneously activated (see also Sect. 3.3.1). First, if someone is unfamiliar to us, mindreading is the key strategy to understand the other and to gain knowledge concerning their individual mindset. This enables us in future interactions to activate top-down those person-models for a rich social understanding which can be fruitful in many everyday situations. Second, we rely on mindreading-person understanding in all those cases where we deliberate about the mental states of another person that we do not directly interact with, for example, in a conversation with a third party about such person or when we think about how our actions might affect them.

To sum up: we employ the ability of top-down activation of person models in combination with the strategy of mindreading in central situations of everyday life. Even looking closely at bottom-up activation makes clear that explicit mindreading is extremely important as some of our prototypical interactions including even simple greeting situations with a person who is significant to us might go beyond direct perception and rule projection and require a flexibility only afforded by mindreading. With our account we oppose the impression a narrow reading of the bottom-up activation has invited for some authors (Fiebich et al., 2016; Gallagher & Hutto, 2008), namely that mindreading and activating an individual mindset is only used as a back-up strategy and that it is also not of special relevance for social understanding.

A few more words shall be said concerning the processes of how the advantages and disadvantages of the respective types of social understanding are balanced. Social understanding has to fulfill two constraints: it has to be cognitively efficient and deliver an adequate social evaluation. Mindreading is the perfect tool to deliver a good social evaluation of an individual which is highly informative but is complex, costly [slow processing according to Kahneman (2011)] and, thus, not cognitively efficient. The other types of social understanding are systematically less complex and less costly [fast processing according to Kahneman (2011)] and, thus, cognitively more efficient but less informative about the individual. Consequently, the most efficient epistemic strategies are at the lowest level of social understanding (Fiebich, 2021). In contrast, mindreading-person understanding is a central tool of social understanding which delivers highly informative social evaluations of individuals but is cognitively costly. It is nevertheless intensely used in daily life because adequate social understanding of individuals is the key ability for success in human societies. Since biological systems

[17] Activating a mindset is often combined with some updating of this mindset in a certain situation. Updating is a more sophisticated form of mindreading than mere activation of a mindset.

have the pressure to be cognitively efficient in each situation the system has to find the ideal balance between using the cognitively most efficient and the cognitively most informative strategy of social understanding. Thus, we call our account balanced pluralism which is plausibly realized by a Bayesian weighting process.

Finally, let us prevent further misinterpretations of our account: We are aware that there might be individual variance in social interest depending on personality traits like introversion/extroversion or cultural differences, for example between predominantly individualist and collectivist cultures. These could lead to strong variations of the dominant use of the types of social understanding. Nonetheless, even across these variations, social understanding remains important and, in the absence of disorders, mindreading abilities are acquired, even if they may be used to varying degrees. Thus, difference in personality and culture modulate our thesis but does not speak against it. It should be noted, however, that our thesis is oriented primarily at Western cultures with an average or strong individualistic orientation, although it may also hold true for many other cultures.

5 Conclusion

Our aim in this paper was twofold. First, we argued for the importance of mindreading within pluralism, which is often underestimated. Second, we aimed to develop an adequate framework describing different types of social understanding and explaining when they are activated and how they interact in order to enable a flexible understanding of others. We offered a systematic overview of the plurality of types of social understanding which result from the interaction of epistemic strategies and models of background knowledge. Which type of social understanding is activated is determined by a Bayesian weighting process relying on bottom-up or top-down activation. In both versions of activation, mindreading plays a crucial role and we demonstrated that it is involved in paradigmatic cases of everyday life. Furthermore, it opens up a new dimension of social understanding, for example, by enabling us to adequately deal with new persons and to account for different mindsets. Finally, we highlighted that types of social understanding integrating mindreading are central in everyday situations even when bottom-up activations are dominant since social understanding is so relevant for us that we not only aim for smooth interaction but also use interaction to learn more about the mindset of a person while the latter of course improves future smooth interactions. Thus, mindreading is not only a back-up strategy but part and parcel of our social understanding.

Funding Open Access funding enabled and organized by Projekt DEAL. We would like to thank the German Research Foundation DFG which supported this research in the context of funding the Research Training Group "Situated Cognition" (GRK 2185/1) and by DFG-project (NE 576/14-1) "The structure and development of understanding actions and reasons". (Gefördert durch die Deutsche Forschungsgemeinschaft (DFG) - Projektnummer GRK 2185/1 und NE 576/14-1).

References

Amadó, A., Serrat, E., & Vallès-Majoral, E. (2016). The role of executive functions in social cognition among children with down syndrome: Relationship patterns. *Frontiers in Psychology, 7*, 1363.

Andrews, K. (2012). *Do apes read minds? Towards a new folk psychology*. MIT Press.

Andrews, K. (2017). Pluralistic folk psychology in humans and other apes. In J. Kiverstein (Ed.), *The Routledge handbook of philosophy of the social mind* (pp. 117–138). Routledge.

Baron-Cohen, S. (1995). *Mindblindness: An essay on autism and theory of mind*. MIT Press.

Brown, E. C., & Brüne, M. (2012). The role of prediction in social neuroscience. *Frontiers in Human Neuroscience, 6*, 147.

Cebula, K. R., Moore, D. G., & Wishart, J. G. (2010). Social cognition in children with Down's syndrome: Challenges to research and theory building. *Journal of Intellectual Disability Research, 54*(2), 113–134.

Coninx, S., & Newen, A. (2018). Theories of understanding others: The need for a new account and the guiding role of the person model theory. *Belgrade Philosophical Annual, 31,* 127–153.

de Bruin, L., & Strijbos, D. (2015). Direct social perception, mindreading and Bayesian predictive coding. *Consciousness and Cognition, 26*, 565–570.

Decety, J., & Sommerville, J. A. (2003). Shared representations between self and other: A social cognitive neuroscience view. *Trends in Cognitive Sciences, 7*(12), 527–533.

Di Paolo, E., & De Jaegher, H. (2012). The interactive brain hypothesis. *Frontiers in Human Neuroscience, 6*, 163.

Fiebich, A. (2015). Varieties of social understanding. *Mentis*.

Fidler, D. J., Most, D. E., Booth-LaForce, C., & Kelly, J. F. (2008). Emerging social strengths in young children with down syndrome. *Infants and Young Children, 21*(3), 207–220.

Fiebich, A. (2021). In defense of pluralist theory. *Synthese, 198*, 6815–6834.

Fiebich, A., & Coltheart, M. (2015). Various ways to understand other minds: Towards a pluralistic approach to the explanation of social understanding. *Mind and Language, 30*(3), 235–258.

Fiebich, A., Gallagher, S., & Hutto, D. D. (2016). Pluralism, interaction, and the ontogeny of social cognition. In J. Kiverstein (Ed.), *The Routledge handbook of philosophy of the social mind* (pp. 208–221). Routledge.

Friston, K., & Frith, C. (2015). A duet for one. *Consciousness and Cognition, 36*, 390–405.

Gallagher, S. (2001). The practice of mind: Theory, simulation, or interaction? *Journal of Consciousness Studies, 8*(5–7), 83–108.

Gallagher, S. (2008). Direct perception in the intersubjective context. *Consciousness and Cognition, 17*(2), 535–543.

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

Gallagher, S., & Hutto, D. D. (2008). Understanding others through primary interaction and narrative practice. In J. Zlatev, T. Racine, C. Sinha, & E. Itkonen (Eds.), *The shared mind: Perspectives on intersubjectivity* (pp. 17–38). Jon Benjamins.

Giaouri, S., Alevriadou, A., & Tsakiridou, E. (2010). Theory of mind abilities in children with Down syndrome and non-specific intellectual disabilities: An empirical study with some educational implications. *Procedia Social and Behavioral Sciences, 2*(2), 3883–3887.

Goldman, A. I. (1989). Interpretation psychologized. *Mind and Language, 4*(3), 161–185.

Goldman, A. I. (1992). In defense of the simulation theory. *Mind and Language, 7*(1 & 2), 104–119.

Goldman, A. I. (2006). *Simulating minds*. Oxford University Press.

Goldman, K. J., Shulman, C., & Burack, J. A. (2018). Inference from facial expressions among adolescents and young adults with down syndrome. *American Journal on Intellectual and Developmental Disabilities, 123*(4), 344–358.

Gopnik, A., & Wellman, H. M. (1992). Why the child's theory of mind really is a theory. *Mind and Language, 7*(1 & 2), 145–171.
Gordon, R. M. (1986). Folk psychology as simulation. *Mind and Language, 1*(2), 158–171.
Gordon, R. M. (1992). The simulation theory: Objections and misconceptions. *Mind and Language, 7*(1–2), 11–34.
Hutto, D. D. (2008). *Folk psychological narratives: The sociocultural basis of understanding reasons*. MIT Press.
Johnson, S. C. (2000). The recognition of mentalistic agents in infancy. *Trends in Cognitive Sciences, 4*, 22–28.
Kahneman, D. (2011). *Thinking, fast and slow*. Farrar, Straus and Giroux.
Korman, J., & Malle, B. (2016). Grasping for traits or reasons? How people grapple with puzzling social behaviors. *Personality and Social Psychology Bulletin, 42*(11), 1451–1465.
Luckett, T., Powell, S. D., Messer, D. J., Thornton, M. E., & Schulz, J. (2002). Do children with autism who pass false belief tasks understand the mind as active interpreter? *Journal of Autism and Developmental Disorders, 32*, 127–140.
Malle, B. F. (2005). Self-other asymmetries in behavior explanations: Myth and reality. In M. D. Alicke, D. A. Dunning, & J. I. Krueger (Eds.), *The self in social judgment* (pp. 109–130). Psychology Press.
Malle, B. F., Knobe, J., & Nelson, S. E. (2007). Actor–observer asymmetries in explanations of behavior: New answers to an old question. *Journal of Personality and Social Psychology, 93*, 491–514.
Newen, A. (2015). Understanding others. In: T. Metzinger & J. M. Windt (Eds.), Open MIND (Vol. 26, pp. 1–28). https://doi.org/10.15502/9783958570320)
Newen, A. (2017). Defending the liberal-content view of perceptual experience: Direct social perception of emotions and person impressions. *Synthese, 194*(3), 761–785.
Newen, A. (2018). The person model theory and the question of situatedness of social understanding. In A. Newen, L. De Bruin, & S. Gallagher (Eds.), *The Oxford handbook of 4E cognition* (pp. 469–492). Oxford University Press.
Newen, A., Welpinghus, A., & Juckel, G. (2015). Emotion recognition as pattern recognition: The relevance of perception. *Mind & Language*, *30*(2), 187–208.
Newen, A. & Wolf, J. (2020). The situational mental file account of the false belief tasks: A new solution of the paradox of false belief understanding. *Review of Philosophy and Psychology*, *11*, 717–744.
Onishi, K. H., & Baillargeon, R. (2005). Do 15-month-old infants understand false beliefs? *Science, 308*(5719), 255–258.
Overgaard, S., & Michael, J. (2015). The interactive turn in social cognition research: A critique. *Philosophical Psychology, 28*(2), 160–183.
Perner, J., & Wimmer, H. (1985). "John thinks that Mary thinks that…" attribution of second-order beliefs by 5- to 10-year-old children. *Journal of Experimental Child Psychology, 39*(3), 437–471.
Peters, U. (2021). Teleology and mentalizing in the explanation of action. *Synthese, 198*, 2941–2957.
Porter, M. A., Coltheart, M., & Langdon, R. (2007). The neuropsychological basis of hypersociability in Williams and down syndrome. *Neuropsychologia, 45*(12), 2839–2849.
Rakoczy, H., Warneken, F., & Tomasello, M. (2008). The sources of normativity: Young children's awareness of the normative structure of games. *Developmental Psychology, 44*(3), 875–881.
Reddy, V. (2008). *How infants know minds*. Harvard University Press.
Ruffman, T., & Perner, J. (2005). Do infants really understand false belief? *Trends in Cognitive Sciences, 9*(10), 462–463.
Schilbach, L., Timmermans, B., Reddy, V., Costall, A., Bente, G., Schlicht, T., & Vogeley, K. (2013). Toward a second-person neuroscience. *The Behavioral and Brain Sciences, 36*, 393–462.
Schuwerk, T., Vuori, M., & Sodian, B. (2015). Implicit and explicit theory of mind reasoning in autism spectrum disorders: The impact of experience. *Autism, 19*(4), 459–468.
Southgate, V., Senju, A., & Csibra, G. (2007). Action anticipation through attribution of false belief by 2-year-olds. *Psychological Science, 18*(7), 587–592.
Spaulding, S. (2015). On direct social perception. *Consciousness and Cognition, 3*, 472–482.
Spaulding, S. (2018). Mindreading beyond belief: A more comprehensive conception of how we understand others. *Philosophy Compass, 13*(11), e12526.
Synofzik, M., Vosgerau, G., & Newen, A. (2008a). Beyond the comparator model: A multifactorial two-step account of agency. *Consciousness and Cognition, 17*, 219–239.

Synofzik, M., Vosgerau, G., & Newen, A. (2008b). I move, therefore I am: A new theoretical framework to investigate agency and ownership. *Consciousness and Cognition, 17*, 411–424.

Thirion-Marissiaux, A. F., & Nader-Grosbois, N. (2008). Theory of mind "beliefs", developmental characteristics and social understanding in children and adolescents with intellectual disabilities. *Research in Developmental Disabilities, 29*(6), 547–566.

Weinberg, M. K., Beeghly, M., Olson, K. L., & Tronick, E. (2008). A still-face paradigm for young children: 2½ year-olds' reactions to maternal unavailability during the still-face. *Journal of Developmental Processes, 3*(1), 4–22.

Wellman, H. M., Cross, D., & Watson, J. (2001). Meta-analysis of theory-of-mind development: The truth about false belief. *Child Development, 72*(3), 655–684.

Westra, E. (2019a). Character and theory of mind: An integrative approach. *Philosophical Studies, 175*(4), 1217–1241.

Westra, E. (2019b). Stereotypes, theory of mind, and the action-prediction hierarchy. *Synthese, 196*(7), 2821–2846.

Westra, E. (2020). When is mindreading accurate? A commentary on Shannon Spaulding's 'How we understand others: Philosophy and social cognition.' *Philosophical Psychology, 33*(6), 868–882.

Wimmer, H., & Perner, J. (1983). Beliefs about beliefs: Representation and constraining function of wrong beliefs in young children's understanding of deception. *Cognition, 13*, 103–128.

Wishart, J. G. (2007). Socio-cognitive understanding: A strength or weakness in Down's Syndrome? *Journal of Intellectual Disability Research, 51*(12), 996–1005.

Wolf, J., & Coninx, S. (2021). The role of mindreading in a pluralist framework of social cognition. *Proceedings of theAnnual Meeting of the Cognitive Science Society*, 43. Retrieved from https://escholarship.org/uc/item/0sp272z1

Zahavi, D. (2011). Empathy and direct social perception: A phenomenological proposal. *Review of Philosophy and Psychology, 2*(3), 541–558.