No facts without perspectives

Author 1—corresponding author:

Dr. Ramiro Glauer

    University of Applied Sciences Potsdam, Germany
    ORCID: 0000-0002-9548-716X

Author 2:

Prof. Dr. Frauke Hildebrandt

    University of Applied Sciences Potsdam, Germany
    ORCID: 0000-0003-3217-6222

Authors have made an equal contribution to the article and appear in alphabetical order.

Abstract

Perner and Roessler (2010) hold that children who do not yet have an understanding of subjective perspectives, i.e., mental states, explain actions by appealing to objective facts. In this paper, we criticize this view. We argue that in order to understand objective facts, subjects need to understand perspectives. By analysing basic fact-expressing assertions, we show that subjects cannot refer to facts if they do not understand two types of perspectivity, namely, spatial and doxastic perspectivity. To avoid conceptual confusion regarding different ways of referring to facts, we distinguish between reference to facts *de re* and *de dicto*.

Keywords

theory of mind, reference, objectivity, subjective perspectives, facts

No facts without perspectives

1.   Introduction

Perner, Roessler, and collaborators have presented a developmental account of how young children understand intentional action that has gained considerable traction over the last few years. In particular, Perner and colleagues hold that children who do not yet have an understanding of mental states explain actions in terms of objective facts (e.g., Perner & Roessler, 2010; Perner & Esken, 2015; Perner et al., 2018). Such explanations of actions are called teleological, and subjects employing such explanations are called teleologists. The claim is that in explaining why, for instance, a subject goes to one box rather than to another to retrieve an object, teleologists assume that the subject goes to that box partly because of the fact that the object is there. According to Perner and colleagues, children who do not yet have an understanding of mental states assume that the objective fact that the object is where it is is part of explaining that action. More generally, these children "see [only] objective facts as providing the reasons for action" (Priewasser et al., 2018, 71, also cf. Perner & Roessler, 2010, 203, Perner & Roessler, 2012, 521). Objective facts play a unique role for the teleological account because "[f]rom the perspective of deliberation, only true propositions—facts—can provide genuine reasons. […] Young children find intentional actions intelligible in terms of 'objective' practical reasons [i.e., facts]" (Perner & Roessler, 2010, 203). 3-year-olds make sense of what one is doing "simply in terms of the worldly facts that constitute good reasons for your action, with no regard to your perspective on your reasons" (Perner et al., 2018, 100). "[C]hildren find actions intelligible in terms of fully objective reasons, relativized neither to the agent's instrumental beliefs nor to her pro-attitudes" (Perner & Roessler, 2010, 205). Even more clearly: "the developmental suggestion we will be pursuing may be put by saying that we all started life as external reasons theorists. For, the suggestion is that young children are familiar with objective reasons before they even grasp that there are two sorts of perspectives from which to consider what someone has reason to do" (Perner & Roessler, 2010, 10).

Perner and colleagues maintain that even older children (and adults) explain the behaviour of others by employing teleology. Their teleology, however, takes others' perspectives into account. It is "teleology within [*S*'s] perspective." Older children interpret S as acting based on what from her perspective appears to be an objective reason (Roessler and Perner 2013: 46; Perner et al. 2018: 100). The transition from the use of "pure teleology," which is the ability to explain behaviour by appeal to objective reasons alone, to the use of "teleology-in-perspective," which involves appealing to an agent's subjective reason(s), is thought to happen at around age 4 (ibid). Thereby,

Perner et al. assume that thinking with objective facts can be conceptually distinguished from thinking with subjective perspectives, i.e., mental states, and that the first precedes the second.

In this paper, we aim to show that children who do not yet have an understanding of subjective perspectives, i.e., of mental states, cannot use the objective fact that an object is in a particular box to explain why a subject goes to that box to find the object. We are going to argue that this is because being able to explain someone's behaviour in terms of objective facts requires an understanding of several contrasts. Most importantly, objectivity cannot be understood other than that which is not merely subjective. This general contrast reappears in several more specific contrasts. Notably, the ability to think in terms of facts at all requires being able to consider states of affairs (SoAs) and to distinguish those SoAs that are facts from those that are not. Moreover, it requires being able to consider these SoAs from different spatial perspectives. As we will see, facts are objective. Therefore, talk of objective facts is a pleonasm and, correspondingly, understanding facts at all involves understanding the contrast between objectivity and subjectivity.

As will be argued, being able to understand these contrasts does not require that subjects know that they act according to those contrasts and that they are thereby able to make these contrasts explicit. All that is required is that subjects are disposed to behave in ways that show that they can make the required distinctions. The aim is to show that being able to explain intentional actions in terms of facts requires an *implicit* understanding of subjective perspectives.

At the outset, we are going to argue, in Section 2, that psychological explanations should capture the 'inner workings' of the cognitive systems in question. That is, the employed terminology should describe how the subjects in question structure their environments. Using the terms we as neurotypical human adults use to structure our environment runs the risk of attributing distinctions and inferences not made by these subjects. In section 3, we are then going to introduce two different notions of reference to facts, namely, reference *de re* and reference *de dicto*. In Section 4, it is argued that the *de re*/*de dicto* distinction concerning reference to facts is not to be confused with the implicit/explicit distinction concerning knowledge or understanding. The latter distinction can only be made once *de dicto* reference to some entity of interest is attributed. Reference to facts *de dicto* does not require an explicit understanding of factuality. In Section 5 and 6, it is then argued that Perner et al. must be committed to attributing reference to facts *de dicto* to young children. In Sections 7–10, we are going to present our main argument that reference to facts *de dicto* involves distinguishing different spatial as well as doxastic perspectives. Together, these two kinds of perspectives comprise understanding mental perspectives. For finding out what is

required for reference to facts, we are going to analyze basic fact-stating assertions. Section 11 concludes.

2. Children's ontologies

A central aim of the debate around young children's capacity to explain intentional behaviour is to elucidate the 'inner workings' of children's minds when engaging in such explanations. These 'inner workings' centrally involve how the environment is structured by young children—which we will call their ontology. Generally, a cognitive system's ontology concerns how the cognitive system structures its environment. An ontology—as we use the term—is not just a list of objects or sets of objects. In this sense, it is definitional that a cognitive system that is incapable of propositional thought has a different ontology than we do.

While there are many behavioural similarities, it is well documented that infants behave differently than adults and, seen from the perspective of an adult observer, often in surprising ways: For instance, infants commit to the A-not-B error (e.g. Piaget, 1963; Smith & Thelen, 2003), infants make errors in individuation tasks (e.g. Stavans et al., 2019), and exhibit peculiar patterns of success and failure in numerosity estimations (Burr et al., 2010). These behavioural differences at least prima facie suggest that infants and young children structure their environment differently than human adults. Furthermore, there are theoretical considerations to the effect that structuring one's environment the way we do is cognitively demanding. We structure our environment in terms of states of affairs which consist of objects that may or may not have certain properties. This involves individuating and classifying objects. Both individuation and classification involve sophisticated cognitive capacities.[1]

From the outset, it is thus an open question, how young children understand intentional behaviour. As Hirsch (1997) lays out, there are indefinitely many ways of structuring one's environment, such as to coordinate one's behaviour with it successfully, i.e. different ontologies. Moreover, many different ways of structuring one's environment can have the same behavioural effects over a wide range of conditions. Therefore similarities in behaviour do not indicate that the environment is

---

[1] For the complications involved in individuation, see the philosophical debate about individual reference (e.g. Russell 1905, Quine, 1960, Castañeda 1966, 1968, and especially Evans, 1982, and Tugendhat, 2016) as well as theoretical reflections on empirical findings on object individuation (e.g. Bermúdez, 2007, Burge, 2010, Butterfill, 2020); for how demanding classification is, see the psychological and philosophical debates about classification and concepts (e.g. Peacocke, 1995, Carey, 2009, Machery, 2009).

4

similarly structured. Only differences in behaviour show that the environment is differently structured. We will shortly consider examples for illustration.

Many ideas of how an organism could see the world differently from us can be found in the philosophical literature. Hirsch (1997) suggests that the observed subjects might, for example, perceive certain discontinuous spacetime portions of reality (Quine 1960). Alternatively, they might structure the world in momentary events and their succession (Hume, 1978). Or, within a Strawsonian early childhood ontology, the perceived environment could consist of placed features and feature changes (Strawson, 1959). All these suggestions have in common that the environment is not being structured into objects and their properties. 'Quineans', 'Humeans', or 'Strawsonians' think of their environment quite differently than we do.

Consider the following illustration. Bats perceive the world quite differently from us. As their primary sense for orientation, they emit high-frequency sounds and process their reflections in order to detect obstacles, food sources, or conspecifics in their environment. Their acoustic sensory input is very different from what our eyes receive and facilitates structuring the environment in quite different ways. As Nagel (1974) famously argued, we will never know what it is like to be a bat.

Nonetheless, by investigating bats' perceptual apparatus and how perceptual inputs are processed, we can at least learn something about the structural features of the environment that most likely play an essential role for bats. The auditory system of some bat species, for instance, uses the Doppler effect to detect objects that move towards or away from the bat. For our brain, such relative motion signals are more difficult to come by and must be processed from the change of depth information, which is a more holistic feature of the visual impression.

This is not to say that bats must have a different ontology than we do. As it stands, this example serves to illustrate the idea that other organisms might perceive their environment in very different ways. That bats see things entirely different than we do is due to significant differences in their perceptual apparatus. However, even if perception is similar, the environment may be structured quite differently. Let us consider an example in which we find it relatively easy to switch between two ways of structuring what we see. Conway's Game of Life consists of a plane grid whose cells can be in one of two states (on/off, live/dead) and which passes through discrete time steps. State changes from time step to time step follow three simple rules: (i) Any live cell with two or three live neighbours survives. (ii) Any dead cell with three live neighbours becomes a live cell. (iii) All other cells die or remain dead in the next time step.
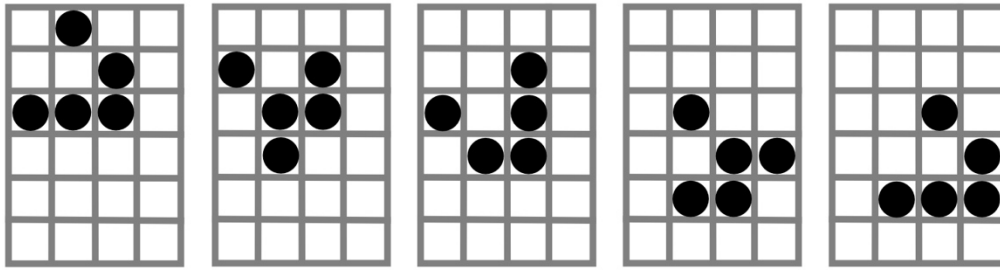
Figure 1: a five-generation cycle of a glider in Conway's Game of Life

In the Game of Life, it is possible to build configurations that recreate themselves cyclically but displaced, within a few time steps (see Fig.1). In visual simulations of the plane, such constellations appear to be moving. Seen under the light of the constituents of the plane, i.e., the cells, however, there is no motion. The cells of the plane always remain where they are. Cells are activated or not; they do not move. There is no quantity or substance that 'spreads' from cell to cell. Expectations about the dynamics of the game can be formed under either description—when thought of as a grid plane of stationary cells turning on and off as well as when thought of as a plane inhabited by moving 'gliders'. This example serves to illustrate the idea that the same sensory impression might be processed in different ways. In the Life-World, we find it easy to switch between the two ontologies: static dots or moving gliders and their kin. This is not to say that any of these ontologies is the correct ontology for the Life-World.

In reality, it could well be the case that other intelligent organisms structure their environment quite differently from us - even if their perceptual systems are largely similar to ours. Thus we may ask: What are the kinds of entities into which the environment is structured for the organism? Moreover, what kinds of regularities do these obey? These could be discontinuous spacetime portions, momentary events and their succession, placed features, objects and their properties, or other entities. How are we to describe these kinds of things such that the structure of the 'inner workings' is best captured?

One particular danger lies in using terminology that facilitates inferences about these 'inner workings' that are not warranted. When describing the behaviour of a Strawsonian feature placer in so-called object individuation tasks, for instance, using object-cum-property terminology may lead to the impression that feature placers commit 'catastrophic individuation failures' which need to be explained.[2] From the perspective of an object-and-properties ontology, it is surprising that someone who can solve individuation problems in certain conditions—and thus appears to understand

---

[2] This terminology is used by Stavans et al. (2019) who use it to describe infant performance in certain individuation tasks. We are not arguing here that infants are Strawsonian feature placers. The example only serves to illustrate that using vocabulary that is suitable for the description of one ontology might hinder understanding the 'inner workings' of a being that comes equipped with a different ontology.

what objects are—is blatantly unable to solve other individuation problems that are just as obvious for us. However, the behaviour that is interpreted as a surprising failure of an otherwise available capacity would not be surprising if performance in individuation tasks would depend on feature-specific expectations about feature changes. In some cases, such expectations would conform to the expectations of someone who individuates objects. In other cases, behaviour diverges.

Young children might structure their environment in a way that does not contain objects and their properties. There is indeed a wide range of criticisms—on various grounds—against the view that very young children individuate objects (Haith, 1998; Cohen et al., 2002; Krøjgaard et al., 2013; Hildebrandt et al., 2020a, 2020b). Moreover, if we are to understand how young children explain intentional behaviour—or any other aspect of their environment—then we must find ways to describe these aspects of the environment in a way that conforms to how they structure their environment. That is, we should determine what kinds of things these children refer to when engaging in explanations of intentional behaviour.

The idea that young children might structure the world differently than adult thinkers is not a far-fetched possibility. It is uncontroversial that children's thinking is under development. And most, if not all, researchers assume that essential aspects of adults' ontologies are not yet in place. Researchers, for instance, investigate how children develop an understanding of abstract categories, numbers, possibilities, subjective perspectives, false belief, or normativity. Trying to understand how object-based thinking develops is just another such question which can be answered in different ways (cf., e.g., Bermúdez, 2007, Burge, 2010).

3.   Notions of reference

We use 'reference' technically rather than in a way it is used in everyday language. Specifically, we take it that reference to X need not involve more than systematically responding or reacting to X.[3] Two different types of reference can then be distinguished, which we shall call, respectively, reference *de re* and reference *de dicto*. In standard usage, the *de re / de dicto* distinction can be applied to any sentence containing an intensional context, i.e., to any sentence whose truth value is not solely determined by the truth values of its component sentences (cf. Nelson, 2019; Garson, 2013). A standard way of illustrating the distinction would attribute a belief to someone in a way that the person herself would not choose to express it. Consider the following:

(S) Lois Lane believes that Superman cannot fly.

---

[3] Note that we are not using 'reference' in the way it is commonly used in the philosophy of language either, where it denotes the semantic relation between singular terms and individual objects.

As Superman is Clark Kent and Lois Lane believes that Clark Kent cannot fly, the belief attribution is, in a way, correct: Lois Lane does believe of the person to which we can refer as Superman that that person cannot fly. This is the *de re* reading of (S). However, if we asked her, she would deny that she would believe any such thing. She knows that Superman can fly. We would have to refer to the same person as Clark Kent in order to elicit her consent. Thus, in another way, the attribution (S) is not correct: Lois Lane would not concede that Superman cannot fly. This is its *de dicto* interpretation. It takes into account how Lois Lane conceives of Superman/Clark Kent. For Lois Lane, the person that is known to her as Superman – the superhero in the red-and-blue spandex costume – can fly. However, the person that she knows as Clark Kent – her modest colleague reporter - cannot (cf., e.g., Nelson, 2019).

The *de re* / *de dicto* distinction is usually used for making explicit how an object can be picked out in different ways from different perspectives. Speaking of *de re* and *de dicto reference* is intended to expand this distinction to cover perspectives from which the world need not be conceived as consisting of objects and their properties at all. *De re* and *de dicto* reference concerns subjects' ways of structuring their environment, i.e., their ontologies. Attributing *de dicto* reference to an aspect of the environment implies that the subject of an attribution structures its environment in the same way as is expressed in the attribution. *De re* reference, on the other hand, merely involves sensitivity to aspects of the environment which *we* see as is expressed in the attribution. It does not involve a commitment to attributing any ontology. Let us illustrate the distinction by way of examples.

Thermostats, birds, dogs, and young children can refer (in the broad sense we assume here) to what we adults call 'facts', but let us assume that they do not have the same ontology as we do. Thus, let us assume that thermostats, birds, dogs, and young children do not structure their environment into objects and their properties and do not think of objects' having properties as facts. Thermostats, for example, may refer to what we would call the fact that it is 35 degrees celsius in the room by flicking a switch. Birds may refer to what we would call the fact that it is 35 degrees celsius in the room by flying to a colder place. Dogs may refer to what we would call 'the fact that it is 35 degrees celsius' by taking a nap. Children may refer to what we would call 'the fact that it is 35 degrees celsius' by saying 'Hot, hot!' All that is required is that an organism or system be disposed to react to what we would describe as a fact. Call this 'reference to facts *de re*'.

Lie detectors, birds, dogs, and small children can refer to what we adults construe as *mental states* as well, by doing certain things when facing someone with what we would describe as a specific mental state. Lie detectors, for example, may refer to what we would call someone's intention to lie by drawing a curve. Birds refer to what we

would call someone's intention to scare them away by flying to a safer place. Dogs refer to what we would call someone's intention to play by wagging their tails. Children refer to what we would call someone's intention to show them a ball by looking at it. Again, all that is required is that an organism or system be disposed to react to what we would call a fact involving a respective mental state. Call this 'reference to mental states *de re*'.

Attributions of reference *de re* are not explanatory by themselves of *how* someone comes to behave in a certain way. They give descriptions of behavioural regularities in a vocabulary that the attributor uses to structure her environment. Attributions of *de re* reference can thus be useful. They can systematize observed behaviour and predictions may rely on them. But they are not meant to capture the 'inner workings' of a system that is thus described. In particular, they are not meant to describe how the system structures its environment or whether it is an intentional system at all.

Whenever some being is not only able to refer to what we adults call facts or mental states *de re* (as, e.g., thermostats, flies, dogs, and very young children do), but also has the same ontology as us, the involved kind of reference to worldly facts or facts involving mental states is what we call 'reference to facts/to mental states *de dicto*'. Again: Having the same ontology is a matter of basically operating in the same conceptual setting. And this, in turn, is a matter of making, on the whole, the same distinctions, operating with the same underlying rules. Most centrally, it involves thinking propositionally, i.e., thinking of the world as consisting of objects and their properties, of propositions as being true or false, of mental states as propositional attitudes. Notice that we do not commit ourselves to the view that thermostats, flies, dogs, and young children can refer to facts or mental states *de dicto*.
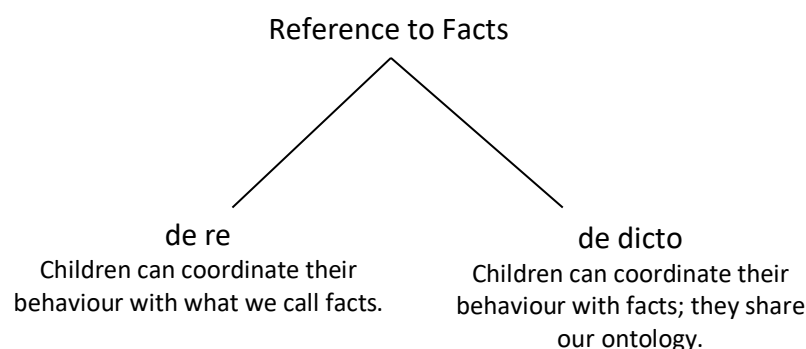
Reference to Facts

```
           /\
          /  \
         /    \
        /      \
```

de re
Children can coordinate their
behaviour with what we call facts.

de dicto
Children can coordinate their
behaviour with facts; they share
our ontology.

Figure 2: Types of reference - *de re*/*de dicto*

Perner et al. propose that young children explain actions by referring to specific facts, namely those facts that would provide objective reasons for actions. With this distinction at hand, we can now ask, what kind of reference to facts Perner et al. have in mind when claiming that children without a grasp of perspectives understand and

explain actions by reference to facts only, not mental states (Perner and Roessler 2012: 521f; Perner et al. 2018: 106).

4.   Reference to objective facts de dicto

It was noted that in order to understand how young children explain intentional behaviour, we should find a description of what children refer to when engaging in such explanations that mirrors the way these children structure their environment. That is, per default, young children should refer *de dicto* to the entities that figure in explanations of their cognitive capacities. Moreover, Perner et al. themselves claim to pertain to what is the case from within the children's point of view. According to Perner et al., 3-year-old and younger children "make sense of what [e.g.] you are doing simply in terms of the worldly facts that constitute good reasons for your action, with no regard to your perspective on your reasons" (Perner et al. 2018: 100). Thus, let us begin by assuming that Perner et al. claim that young children refer to objective facts *de dicto* but do not refer *de dicto* to facts involving mental states. This implies that objective facts are part of the ontology of young children, but mental states are not. For a subject S to refer to facts *de dicto*, S needs to think of her environment as being structured into objective facts. That is, S needs to share (this aspect of) the ontology of adults.

For our present purposes, all that is required is a minimalist account of shared ontology regarding facts and mental states. According to this account, for S to share an ontology with an adult means for S to share what we call an implicit understanding of facts and mental states with adults. For S to have such an understanding, in turn, means for S to be able to use the concepts involved in thinking about facts or mental states according to the rules that would specify the use of these concepts for us. That is, S's implicit understanding of facts is embodied in S's behavioural dispositions. Explicit understanding is present when the rules governing these dispositions can likewise be accessed (Dennett, 1982, 221; also cf. Perner & Roessler, 2012). Crucially, implicit understanding does not require S also to represent the rules themselves that capture the dispositions; following them is sufficient. To illustrate: Being able to speak a language requires being able to follow a vast number of grammatical rules. It does not also require having access to these rules. If S also represents these rules themselves and is, therefore, able to express them, she has an explicit understanding of facts and mental states.

Note, thus, that the *de re*/*de dicto* distinction does not coincide with the implicit/explicit distinction. Applying the implicit/explicit distinction presupposes that subjects are already taken to follow the rules that are characteristic of a given concept. Therefore, it only applies to attributions of *de dicto* reference to worldly facts or facts

involving mental states. To repeat: we are not claiming that reference to facts *de dicto* requires an explicit understanding of factuality or mental states.
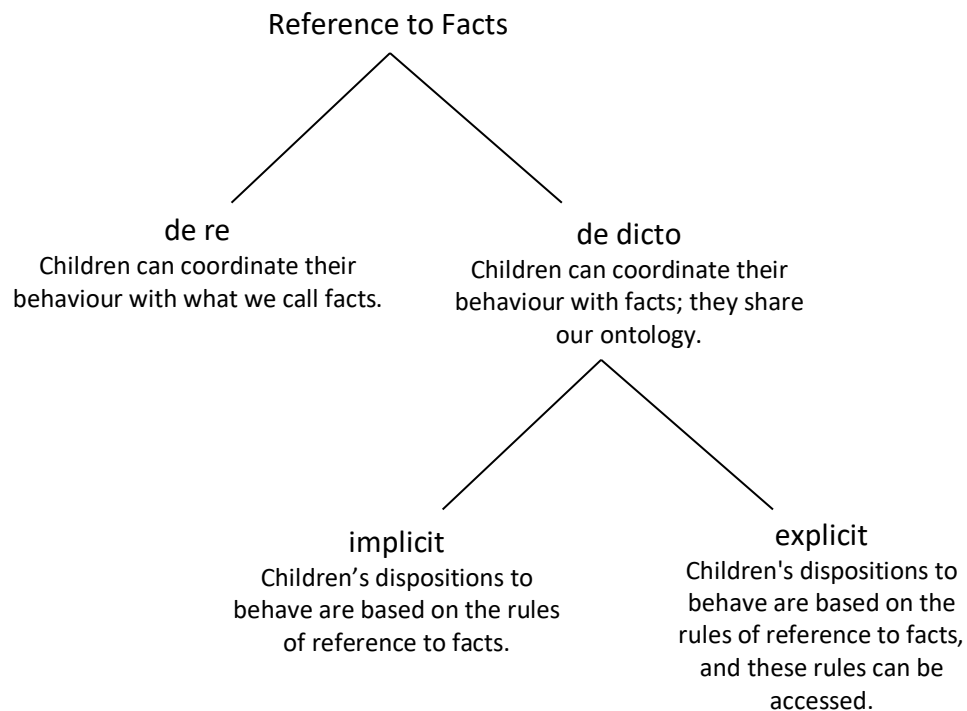
Reference to Facts

de re
Children can coordinate their behaviour with what we call facts.

de dicto
Children can coordinate their behaviour with facts; they share our ontology.

implicit
Children's dispositions to behave are based on the rules of reference to facts.

explicit
Children's dispositions to behave are based on the rules of reference to facts, and these rules can be accessed.

Figure 3: Types of reference - *de re*/*de dicto* and implicit/explicit

Holding these two distinctions apart lets us see that the formulation 'understanding X as X' is ambiguous. It may either appeal to the *de re*/*de dicto* distinction, in which case it means that another subject does not merely react to something we would call 'X' but operates based on the same rule as we do when employing 'X'. Alternatively, saying that S understands X as X can mean that S is not merely able to follow the rule but is also able to represent and therefore express the rule that is operative in applying 'X'. Thus used, the 'as'-formulation appeals to the implicit/explicit distinction.

The "X as X"-ambiguity is the result of not keeping apart these two very different distinctions. When relying on the *de re*/*de dicto* distinction, "understanding X as X" means "understanding X *de dicto* as opposed to understanding X *de re*". When relying on the implicit/explicit distinction, "understanding X as X" means "understanding X *de dicto* explicit as opposed to understanding X *de dicto* implicit". The result of this confusion is that "understanding X as X" can be interpreted as "understanding X *de dicto*" (simpliciter)—which is compatible with an implicit understanding of X *de dicto*. Or it can be interpreted as "understanding X de dicto explicit". In short, we can say that the ambiguity concerns interpreting "X as X"-formulations as being intended *de dicto* (simpliciter) versus being intended *de dicto* explicit.

Arguably, using the 'as'-formulation can become problematic when it leads to conflating *de dicto* ascriptions and explicit understanding as it conceals viable theoretical options. Explicit understanding is second-order (having a concept and having access to it). *De dicto* ascriptions can be first-order (having a concept but not having access to it) as well as second-order (having a concept and having access to it). But this distinction does not even apply to mere *de re* ascriptions, because they are not intended to capture the 'inner workings' of another individual.

Whether infants have an explicit understanding of the rules underlying the use of certain concepts is not the topic of this article. We are concerned with whether young children have an implicit understanding of objective facts. And the attribution of implicit understanding of objective facts is problematic. In brief, on the minimalist reading of sharing an ontology, the concepts employed in an attribution must capture the distinctions made by the subject of an attribution. That is, attributing objective fact understanding to young children must capture the distinctions made by young children. However, as will be argued in the next section, objectivity is a rather demanding notion that contrasts to subjectivity such that the distinctions required for understanding objectivity can only be made in conjunction with those required for subjectivity. Therefore, understanding objectivity *de dicto* without understanding subjectivity *de dicto* is impossible.

5.   The contrast between objectivity and subjectivity

According to the picture promoted by Perner et al., infants first acquire an understanding of objective facts and only later realize that facts are always taken to obtain from a particular perspective. Objectivity, however, can hardly be conceptualized as just-so taking what appears to be the case. Objectivity is commonly characterized in contrast to subjectivity. In everyday usage, objectivity is achieved by refraining from subjective interests or biases (e.g., Objectivity, n.d.). If we were to explain objectivity to someone who does not yet understand the notion, we would have to proceed in the following way:

> "Everyone sees the world from her perspective, from a particular standpoint. Take this apple, for instance, we say that it is the same for every one of us - and even independently of us. Even if no-one of us would be here to see the apple, it would still be there as it is. Thus we say that it is objectively here."

Objectivity is explained in contrast to subjective perspectives. It holds independently of any perspectives. It is not the default position of someone who does not yet understand that different people can have different perspectives. Facts objectively obtain independently of any perspective ('from nowhere'; cf. Nagel 1986). Also, note the close connection between facts and objectivity. Facts are obtaining states

of affairs (SoA). If an SoA obtains, it obtains objectively, i.e., independently of any perspective.

At the same time, subjectivity cannot be understood without understanding objectivity. In order to understand that one always sees things from a particular perspective, one must understand that there is something on which one has a perspective. Different people can have different perspectives only if there is something about which they can disagree, i.e., objective facts. If we were to explain subjectivity, we would have to rely on a notion of objectivity:

> "Take this apple. We say, it is objectively the same apple for all of us - and even independently of us. We can all see it, but it appears different to each of us. We say that we see it from different perspectives, from our subjective standpoint."

In academic contexts, objectivity is likewise characterized in terms of subjectivity. For instance: "[s]cientific objectivity is a characteristic of scientific claims, methods and results. It expresses the idea that the claims, methods and results of science are not, or should not be influenced by particular perspectives, value commitments, community bias or personal interests, to name a few relevant factors" (Reiss & Sprenger, 2017).

Objectivity and subjectivity are mutually interdefined. Speaking of objectivity only makes sense in contrast to how the world might be taken to be from a subjective perspective. Understanding objectivity involves understanding its independence of any perspective: One has to understand what one has to abstain from. Correspondingly, objective facts cannot be understood *de dicto* without understanding subjectivity.

Perner et al.'s understanding of objectivity does not diverge from this characterization. As we have seen, objective facts obtain independently of how and whether anyone thinks of them (cf. Perner & Roessler 2010: 205)). And objective reasons are understood as independently existing facts in the world (cf. Perner et al. 2018: 100). I.e., Perner et al. understand objectivity as independent of anyone's perspective and characterize objectivity in contrast to subjective perspectives. Nonetheless, they hold that teleological reasoning is not tied to understanding perspectives. As a result, Perner et al. cannot claim that young children understand objective facts as we do, which means that they cannot claim that children understand objective facts *de dicto*. But how are we to understand their claim that young children make sense of others' behaviour in terms of objective facts?

That young children can refer to objective facts but not to facts about mental states might be understood as an auxiliary formulation that attempts to capture differences, but also similarities, between the ways adults and young children structure their environments. It would thus be an attempt to find a description of young children's

'inner workings' in their terms. Young children would refer to "objective facts" *de dicto* but not to objective facts. While understanding objective facts requires understanding subjective perspectives, understanding "objective facts" does not.

Calling whatever young children refer to when explaining intentional behaviour "objective facts", however, is unfortunate. As reference to objective facts is attributed without attributing understanding of subjective perspectives, it is clear that not all distinctions and implications carried by the involved notions can be used to describe and explain children's behaviour. However, it is not further specified which aspects of the notions are to be carried over. If "objective fact" does not have its ordinary meaning, what does it mean? Perner et al. owe us a clear definition of what reference to "objective facts" shall amount to. Also, the notion of an objective fact is so closely connected to other aspects of our ontology that it is difficult not to draw unwarranted conclusions. Starting in Section 7 we will come to see that reference to facts *de dicto*—not just reference to objective facts—is indeed so tightly intertwined with understanding perspectives, truth, objectivity, and possibility that it is hard to see what is left of it when an attempt is made at separating it from subjective perspectives. But first, let us consider whether Perner et al. might merely intend to attribute reference *de re* to objective facts to young children.

6.   Reference to objective facts *de re*

From the outset, attributing reference to objective facts *de re* is not very attractive overall. Remember that *de re* attributions do not carry any commitment to how a subject structures her environment. In *de re* attributions, our way of structuring the environment is merely used to describe a behavioural regularity. This description can serve as a starting point for finding a proper explanation of a cognitive ability in the subject's terms. Nevertheless, *de re* attributions are not explanatory by themselves of *how* someone comes to behave in a certain way.

Furthermore, interpreting Perner et al.'s claims as attributing reference to objective facts *de re* bears the additional difficulty that 'understanding' in 'understanding objective facts' and in 'understanding perspectives/mental states' is not used with the same meaning. In denying young children ' understanding perspectives/mental states,' Perner et al. are quite clear that this is meant *de dicto*. For one, there is textual evidence (cf., e.g., Priewasser et al., 2018, 71, Perner & Roessler, 2012, 521, Perner & Esken, 2015, 77–78). For another, denying reference to facts involving mental states *de re* is not very plausible. As we have seen, we have no difficulty ascribing the ability to refer *de re* to both worldly facts and facts involving mental states even to relatively simple organisms and machines. Remember that, in the current terminology, lie detectors can refer *de re* to intentions to lie. Moreover, there is good evidence that young children do

refer *de re* to what for us are mental states such as, for instance, intentions, emotions, desires, and beliefs (e.g., Trevarthen, 1979, Stern, 1985, Tomasello, 2003, Apperly and Butterfill, 2009, see Tomasello, 2019, 308 ff. for an overview). If reference to objective facts were attributed *de re*, their use of 'understanding' would be homonymous. Nothing substantial would follow if Perner et al. were to claim that young children cannot refer to facts involving mental states *de dicto* but can refer to objective facts *de re*.

Thus, if Perner et al. (i) want to make an *explanatory* claim about children's reference to facts and (ii) do not employ a homonymous notion of reference, they cannot attribute reference to facts *de re*. It is therefore fair to say that when Perner et al. claim that young children can initially only refer to objective facts, but not to mental states, they should *not* mean the *de re*-claim. They must claim that young children can initially only refer to objective facts *de dicto*, but not yet to facts involving mental states. We are thus back at the *de dicto* reading of understanding facts that we have rejected on the ground that understanding objective facts presupposes understanding subjectivity.

But maybe Perner et al.'s talk of objective facts was not intended to contrast objectivity and subjectivity. After all, the notion of an objective fact is primarily brought in for distinguishing two kinds of action explanation: one in terms of the beliefs and desires that lead to acting in a certain way and one in terms of the external facts that speak for acting in that way (cf., e.g. Perner & Roessler, 2012). Maybe, Perner et al.'s use of 'objective fact' is merely meant to capture this distinction and the qualification 'objective' is meant to highlight that the facts involved in the envisaged action explanations are 'worldly' or 'external' facts as opposed to facts including 'mental' or 'internal' states. They might think that their account only depends on children's ability to understand which *wordly* facts are relevant for someone's actions, without additionally understanding that these facts are *objective*. However, in Sections 7 to 10, we are going to argue that (implicitly) understanding *facts whatsoever* requires (implicitly) understanding subjective perspectives, which means that it requires understanding objectivity as well. Facts simpliciter cannot be understood *de dicto* without understanding the contrast between subjectivity and objectivity.

To this, it might be objected that the employed notion of "implicitly understanding facts" is still too demanding and that an ontology involving facts might simply be imposed onto a cognitive system by being struck by facts in perception. Thereby, an organism might refer *de dicto* to facts without having to understand the contrast between objectivity and subjectivity. In Sections 11, we are going to argue that even if facts simply strike in perception, the organism still has to be able to (implicitly) make all distinctions that are characteristic of reference *de dicto* to facts.

7.   The contrast zoomed in: Understanding assertions about worldly facts presupposes understanding perspectives.

We shall assume that an implicit understanding of facts *de dicto* becomes manifest in the linguistic expressions adults use to refer to facts *de dicto*, that is, in assertions, and can be made explicit by attending to the rules of their use. The idea that we can find out structural invariants of how we think by investigating certain aspects of language is in line with so-called measurement accounts of propositional attitudes (Field, 1981, Davidson, 2001a, Churchland, 1979, Dennett, 1987, Beckermann, 1996, Matthews, 2007). The central idea is that attributions of propositional attitudes semantically function like measurement statements. In measurement, relations among the objects of a measurement domain (e.g., objects that have a mass for measurement of weight) are represented in an abstract structure—commonly the natural numbers—in a way that preserves the respective relations among the objects of the measurement domain. Similarly, in attributions of propositional attitudes, the structure of propositions preserves relations among certain mental states (the propositional attitudes). By analyzing the structure of propositions, we can, therefore, learn about the structure of these mental states. For this, it is not necessary to assume that reference to facts *de dicto* requires language or that reference to facts *de dicto* is linguistically structured.

Nonetheless, considering the rules that govern the use of linguistic expressions can help us obtain an idea of the conceptual complexity involved in their usage. In assertions, our implicit understanding of facts is expressed. By analyzing the use and meaning of assertions, the distinctions and implications that are characteristic of reference to facts *de dicto* can be brought to the fore. Any creature that can refer to facts *de dicto* must be able to make distinctions that are characteristic of such an ontology and be disposed to behave in corresponding ways.

Assertions are linguistic expressions that serve to make a statement about how the world is. If an assertion is to be true, the state of affairs (SoA) expressed by the assertion must be a fact. Assertions express what a speaker takes to be a fact. Correspondingly, in order to make an assertion, speakers must be able to distinguish what is the case from what is not. That is, they must understand SoAs and distinguish those SoAs that are facts from those that are non-facts (falsehoods).

In order to understand the assertion 'This apple over here is green', for instance, one needs to understand that this over here is *an apple* and *green*. Furthermore, one has to realize that *this over here* is an apple and green. That is, one needs to know the usage rules of the relevant general terms/predicates (in the example sentence, to which colour 'green' applies and which objects are correctly classified by the general term

'apple'). However, in order to understand that this is a green apple, it is not enough to be able to distinguish green things from non-green things and apples from non-apples. In order to structure one's environment into facts, one also has to understand the usage rules of the involved singular terms (how to use 'this over here' for picking out a *particular object*) and how both kinds of terms are combined to form an assertion. We are going to focus on singular terms and introduce two different kinds of understanding of perspectives regarding SoAs in the next section. There is what we call an 'understanding of a *spatial perspective*' and an 'understanding of a *doxastic perspective*.'

8.   Understanding spatial and doxastic perspectives

Perry (2000), Casteñeda (1966; 1968), Kaplan (1977), and others have influentially argued that indexicality plays an essential role in language and thought.[4] Sentences involving indexical expressions are not reducible to sentences without indexicals. However, the importance of indexicality arguably runs even deeper. As Evans (1982) and Tugendhat (2016) have suggested, non-indexical reference indeed depends on indexicality because any perceptually experienced referent (as opposed to abstract ones, which we shall here set aside) is ultimately fixed demonstratively.

'Here' is a spatial indexical, i.e., a context-dependent expression for a location. 'Here' refers to a place but—as opposed to names of locations such as "Sanssouci"—it does not always refer to the same place. The location referred to changes with where the expression is uttered. Roughly, by saying "here," a speaker refers to her own location. Thus, the crucial information someone needs to grasp if she is to understand an utterance of "here" is the position of the speaker. That is, in order to understand what a person refers to when she says "here," one has to know where she is located. For understanding "here" it is crucial that *any speaker* can refer to her actual position by saying "here." Grasping the meaning of "here" requires understanding that what a person refers to by saying "here" is *her own* position.

Moreover, for a person to be able to refer to her own position with 'here,' she must know (implicitly) that she has a particular position among others who have different positions. Locating oneself in a shared, public space requires an implicit conceptualization of oneself as an object in space. And speakers can refer to that object whose position can always be referred to by saying "here" with the expression "I" (see also Tugendhat, 2016, 354 ff.).

---

[4] Note that we are currently concerned with singular reference, that is, reference to particular objects, not with general reference to classes of objects or properties. There are fundamental differences between singular and general reference which should be held apart clearly.

Furthermore, and most importantly for our current interest, in order to be able to use 'here' correctly (*de dicto* / as we do) a speaker S1 has to understand that 'here' for her can be the same place as 'there' for another speaker S2. If S2, standing some meters away from S1, wants to refer demonstratively to the same place to which S1 refers to by saying 'here,' S2 needs to use *another* term to specify the *same* place. Because 'here' always refers to the position of the speaker him- or herself, by using 'here', speaker S2 could only refer to his own current position, not to the (different) current position of S1. To refer to S1's position, S2 can use the term 'there.' 'There' indeed always refers to a position that is different from the speaker's and to which someone else, being in the right place, could refer to by saying 'here.' This systematic inter-definedness between 'there' and 'here' is just part of the usage rule of demonstratives. The same holds, mutatis mutandis, for 'this' and 'that'.

By being able to use "here" adequately, a speaker implicitly knows that she has a position in space among others who have different positions. That is, she has to understand that distinct objects (including other persons) have different positions in space and that different speakers can refer to the *same* place from different perspectives by different terms. She knows that 'here for me' = 'there for S.' This is what we mean by an understanding of *spatial perspectives*.

Spatial perspectivity, here, is not meant to capture the truism that things might look different from different angles. That is undoubtedly a form of spatial perspectivity but one that holds for the usage of general terms. (Does it look green from there? Does it look round from there?) Currently, we are interested in one form of perspectivity that is inherent to the usage rules of certain singular terms. It concerns different aspects of demonstratively picking out the same individual object from different positions in space. The correct usage of 'this' and 'that', 'here' and 'there' requires the ability to make such perspectival shifts between different positions in space. Otherwise, a speaker would not refer to a particular object that is located in shared space.

There is another type of perspectivity that is required to understand the worldly-fact sentence "This apple over here is green." If a speaker S is to understand this assertion, S needs to understand that "This apple over here is green" can be either true or false. Indeed, it is one of the central insights of twentieth-century philosophy of language that the meaning of a sentence is closely tied to its *truth-conditions*. As Wittgenstein (1922) famously put it: "To understand a proposition (einen Satz) means to know what is the case if it is true. (One can therefore understand it without knowing whether it is true or not)" (ibid., 4.024). Notice that understanding an assertion need not include believing that, or knowing whether it is true or not. What is required is grasping what *would* be the case *if* the assertion were true. One can understand an assertion p (a) without having any belief concerning its truth value and (b) without being in a situation to evaluate it.

In other words, understanding p involves being able to imagine in which situations p could be evaluated without having to be in such a situation. Understanding an assertion thus requires considering the SoA expressed by an assertion without, thereby, taking it to be true. SoAs can obtain or not obtain; they are possibilities. If understanding p is possible without believing that p is true (and one might say that understanding p is indeed a precondition for believing that p—because how could one believe that p without understanding the meaning of p), understanding p entails (implicitly) grasping p as not yet evaluated, as possibly true or false, that is, as a state of affairs. "Grasping the difference between facts and states of affairs" is a reformulation of "entertaining a proposition without judging it to be true" or "entertaining it as a possibility". That is, insofar as understanding an assertion implies knowing its truth-conditions, understanding assertions involves an implicit understanding of possibility. This need not, of course, amount to an explicit understanding of possibility that could be expressed by the use of modal operators ("possibly", "necessarily", "actually", ...). It merely involves the ability to entertain a proposition without judging it to be true.

To understand the assertion 'This apple over here is green' one has to be able to grasp what would be the case if the sentence were true, i.e., if the expressed SoA were a fact. For this, S must at least be able to consider the SoA *without* taking it to be the case. Understanding an assertion (by knowing its truth-conditions) is not the same as taking it to be true.

Thus, to understand assertions, S needs to be able to entertain propositions without taking them to be true, and for this, in turn, S needs to implicitly grasp the difference between the two: an SoA and a fact. She must be able to understand (implicitly) that the SoA may or may not obtain.

This epistemic distance facilitates entertaining beliefs (again, implicitly) by considering SoAs. SoAs comprise the contents of beliefs, and beliefs are individuated by their contents. By considering SoAs, one *ipso facto* considers possible beliefs. This does not yet amount to an explicit understanding of belief. It comprises implicitly understanding belief by understanding that any SoA could obtain or not obtain and that any assertion could be true or false. That is why we call it 'an understanding of *doxastic perspectives*.'

Among other things, our distinction of *de re* and *de dicto* reference is intended to clarify in which sense cognitive beings have a perspective. *De re*, any creature that is able to perceive has a perspective simply by being located in spacetime. But in order to have a perspective *de dicto* that could be used to explain how a cognitive being structures its environment, a creature must implicitly understand that it is a perspective on an independent object on which others could have different perspectives (compare

Tomasello's characterization of shared intentionality, 2019, How children come to understand false beliefs). Having a perspective *de dicto* involves distinguishing one's perspective from what it is a perspective on, which means distinguishing subjective from objective.

9.   Doxastic perspective and spatial perspective in interaction

Both the spatial perspective and the doxastic perspective are involved in understanding a simple assertion like 'This apple is green.' In order to understand which object 'this' refers to, speakers must understand that the object is located in shared space and can be referred to differently from different spatial perspectives. Moreover, to understand that the sentence asserts a fact involving that object, speakers must understand that the SoA expressed by the sentence could also be false. From the inter-definedness of 'here' and 'there' it follows that in order to understand that this over here is a green apple, one has to understand, implicitly, that ['this over here (from my spatial perspective) '*equals* 'that over there (from S's spatial perspective)'] is an apple and green.

Note that the interdefinedness of spatial demonstratives concerns their peculiar substitution rules within the context of a sentence. The substitution rules of "here" and "there" and "this" and "that" do not stand freely of what can be said about places or objects. Understanding assertions involving such demonstratives requires being able to understand that "This over here is a green apple," as said by me, has to be expressed by saying "That over there is a green apple" by someone else standing away a few meters.

Thus, our example sentence unpacks into several propositions. And anyone who has a *de dicto* fact understanding needs to consider several propositions and examine whether they are true: A proposition stating that the object, demonstratively referred to from S's perspective by saying "this over here", is a green apple, and an unspecified number of propositions stating that the object, demonstratively referred to from other perspectives by saying "that over there", is a green apple. Correspondingly, at least the following two questions have to be settled:

(1) Is it the case that this over here, from my spatial perspective, is an apple and green?

(2) Is it the case that that over there, from S's spatial perspective, is an apple and green?

In order to be able to ask oneself these questions, one has to understand that 'this, from my spatial perspective, is an apple and green' may or may not obtain, i.e., that it is an SoA. Furthermore, one has to understand that the SoA 'that over there, from S's spatial perspective, is an apple and green' may likewise obtain or not obtain. For both,

as shown above, one must be able to think of the SoA independently of whether it is a fact. In short, one has to understand that the sentence "That over there from S's spatial perspective is an apple and green." can be false.

Notice that this is nothing else than recognizing that someone else *might be mistaken*, i.e., the classical understanding of mental states measured in false-belief tasks, which children begin to master at around age 4–5 (Wellman et al. 2001). An understanding of facts *de dicto* goes along with the ability to grasp that others have different doxastic and spatial perspectives and that there is the possibility of error. This is precisely what counts as an understanding of mental states *de dicto*: "We may say that beliefs are mental representations that their possessor takes to correspond to an objective reality but which everyone who understands such things knows may not." (Tomasello 2018: 8491). Or, similarly: "If false-belief difficulties do reflect a conceptual limitation, then what is the nature of this conceptual deficit? Subjects have to understand that another person will assign a conflicting truth value to a critical proposition (e.g. 'Box contains Smarties' is TRUE) which conflicts with the value they themselves assign (i.e., 'Box contains Smarties' is FALSE)" (Perner, 1987, 135).

One understands the assertion "This is an apple and green" only if one considers whether it is true, not only from one's own spatial perspective but also from S's perspective. And, of course, if one believes that it is a *fact* that this over here is an apple and green, at the same time one believes that it is a fact seen from S's spatial perspective. Thus, an understanding of facts *de dicto* goes along with an understanding of mental states *de dicto*.

Note that in order to understand mental perspectives, a subject must understand that the following two propositions can be evaluated differently: that this over here (from my perspective) is a green apple and that that over there (from S's perspective) is a green apple. By holding-true that this over here (from my perspective) is a green apple one is committed to believing that that over there (from S's perspective) is a green apple. In order to make this commitment, in turn, one must be able to consider the two SoAs separately. Otherwise, one could not be said to think that the sentence "This over here is a green apple" states a fact. If a basic fact-stating sentence is true, it is true from any perspective. As we have seen, that SoAs are to be considered from several perspectives is inherent to the usage rules of spatial demonstratives. It amounts to implicitly understanding that propositions could be true or false (SoAs could obtain or not, could be facts or non-facts) together with understanding that whether a proposition is true (whether an SoA obtains) has to be considered from different spatial perspectives. This is the standard characterization of understanding mental perspectives.

10.   Knowing whether someone has a perspective

The characterization of spatial and doxastic perspectives does not yet consider whether the subject (or thing) which occupies a position in space from which "this" could be referred to by "that" also has a perspective (de dicto). As it stands, the "subject" of spatial and doxastic perspectives could be any object, for instance, a thermostat. A thermostat can be situated at a position from which "this over here (from my perspective)" equals "that over there (from the thermostat's perspective)." While it occupies a spatial position by being located in space, it does not have a perspective, because it does not even implicitly understand that objects can be referred to differently from different positions in space, let alone, that propositions could be true or false. Having a perspective is a matter of implicitly understanding that the same object can be referred to differently from different positions and that propositions could be true or false.

Whether someone (or something) can implicitly make these distinctions is not directly observable. If understanding spatial and doxastic perspectives is to amount to an understanding of mental perspectives, however, the ability to take on such perspectives must involve an ability to detect whether someone (or something) else equally understands these perspectives and thereby *has* a perspective. How can we know whether someone (or something) actually has spatial and doxastic perspectives and does not (1) merely occupy a position in space from which one *could* have a spatial perspective and (2) merely react to facts (de re) that *could* be taken not to be facts? Fortunately, our analysis of individual reference provides linguistic criteria that can serve as a sufficient condition for whether someone (or something) occupies a spatial position and *has* a perspective.

Being able to use "here" *de dicto* is to be able to use "here" in accordance to its usage rules, including, most notably, the pattern of substitutions of "here" and "there." If a person is able to use "here" *de dicto*, as indicated by her conforming to its usage rules, she understands implicitly that by "here" she refers to her own position in space. By using "here" appropriately, a speaker shows that she conceptualizes herself as an object in space having a particular spatial perspective. Thus, that someone conceptualizes herself as an object in space with a particular spatial perspective that differs from others' can be concluded from her ability to use "here" *de dicto*. It is true not only for us but for all subjects who are able to use "here" *de dicto* that they implicitly understand that there are different spatial perspectives on the same object.

Thus, anyone who is to understand the sentence "This over here is an apple and green" (i) has to understand (implicitly) the SoAs "This over here from my spatial perspective is an apple and green" and "That over there from someone else's spatial

perspective is an apple and green" and (ii) has to understand (implicitly) that anyone who utters this sentence *de dicto*, i.e., in accordance to the usage rules, likewise understands both SoAs in the same way. This is because, as argued above, any utterer of such an assertion must (implicitly) understand doxastic and spatial perspectives. Whether someone utters assertions like "This (over here) is a green apple" appropriately can thus be used as a criterion for whether someone has a spatial perspective. Moreover, anyone who can use "here" *de dicto* has to be able to understand that the assertion "This over there (from the perspective of someone else) is an apple and green" can be true or false. That is, she must have a doxastic perspective.

11.   Perceiving facts

Perner et al. might attempt to avoid this conclusion by arguing that worldly facts are simply what we encounter when we perceive while facts involving mental states do not thus present themselves in perception. The underlying view seems to be that we are simply struck in perception by worldly facts. Thereby, the environment imposes an ontology onto our cognitive system. Thus, young children can refer *de dicto* to worldly facts in their explanations of behaviour but not to facts involving mental states[5]

The suggestion that facts are just what we encounter when we perceive, however, is problematic on several counts. First, as noted in Section 2, the environment can be structured in various ways that can all enable an organism to successfully coordinate its behaviour with its environment. Abilities of feature discrimination and feature pattern generalization, for instance, can produce behaviours that are equivalent to adult human performance in a wide range of conditions (Hildebrandt et al., 2020a, 2020b) without structuring the environment into facts. Generally, behavioural similarity alone does not provide evidence for sharing an ontology. In particular, this means that behaving in similar ways as human adults does not provide evidence for reference to facts de dicto.

Second, perceiving facts *de dicto*, that is, sharing an ontology, is more committal than suggested here. It involves a commitment to the psychological relevance of the attributed ontology, including the pattern of inferences into which the terms used to specify the ontology are embedded. In the case of "fact", taking something to be true is definitional. The expressions "it is true that…" and "it is a fact that…" can regularly be used interchangeably.[6] This means that attributing an ontology of facts (de dicto) does

---

[5] This idea was brought up by Josef Perner, Johannes Roessler, and Eva Schmidt in a discussion about an earlier draft of this article as well as by an anonymous reviewer.

[6] also cf. Oxford Dictionary of English: fact | fakt |
noun
• a thing that is known or proved to be true …
• [mass noun] mainly Law the truth about events as opposed to interpretation …

not merely attribute an ability to react discriminately to what we call different facts (de re). Discriminately reacting to different facts de re can be achieved by various cognitive means or ontologies. In order to coordinate one's behaviour with the fact (de re), say, that there is a green apple as opposed to a banana, an organism merely has to be able to discriminate green-apple features from banana features.

Attributing an ontology of facts, in contrast, critically involves attributing the ability to implicitly distinguish between actually being the case (that is, truth) and not being the case (that is, falsity). This requires a sentential or propositional structure. "a is F" is true if and only if the referent of "a" satisfies "is F" (see Tarski, 1935, 1944). Without a propositional structure, a cognitive system would not be following truth-rules (whether a is F) but some other form of correctness rule (something like "'F' if F"). In the latter case, the cognitive system would not refer *de dicto* to facts but merely "detect F-ness". Distinguishing between truth and falsity is achieved by examining whether the thing referred to really is an apple or not. In order to examine whether something is an apple, one has to be able to entertain the proposition whether the thing is an apple independently of taking it to be true or false already. As a result, the default of attributions of mental states about facts involves attributing an ability to distinguish truth from falsity and thereby to entertain propositions without taking them to be true or false already. As argued, this amounts to an implicit understanding of perspectives.

Perner et al. might respond that expressions may be used in a sense that differs from common usage such that Perner et al. might intend to use "fact" in a sense that does not imply truth. Using our ordinary terms, however, is problematic in the present context. Taking a phrase from Glock who uses it in the context of applying mental expressions to non-human animals, attributing fact understanding to very young children is "incongruous in that the rich [...] idiom we employ has conceptual connections that go beyond the phenomena to which it is applied" (Glock, 2000). When not intending to commit to attributing an understanding of truth when attributing facts as mental contents, it must be made clear in which sense the use of "facts" in a given context diverges from common usage. Perner et al. would have to specify the differences of the intended meaning of "fact" if they do not intend to use "fact" with its common meaning. Furthermore, they would have to take precautions not to draw inferences that are based on common usage but not admissible for the restricted usage. Perner et al.'s account of action explanation, however, relies on the full notion of a fact.

Most centrally, (certain) worldly facts are to play the role of reasons for action in teleological explanations. The whole apparatus of reasons, of evidential support of beliefs, and of practical support of actions depends on truth versus falsity and on doing right versus doing wrong (Davidson, 1963). Perner et al. explicitly endorse such a propositional view of rationality and of reasons for action (Roessler & Perner, 2013,

44). In this sense, if facts are to play their role as reasons, (implicitly) understanding facts must involve (implicitly) understanding truth and falsity and thereby perspectives.

As a result, the idea that the world imposes its structure onto an organism by striking it in perception does not alleviate the requirement of choosing carefully the terms with which the organism's ontology is described. Even if facts simply strike in perception, the organism still has to be able to (implicitly) distinguish truth from falsity and to consider propositions without taking them to be true or false. As it is near impossible to avoid inferences that are based on common usage, it is recommendable not to use such loaded terms when intending a heavily restricted sense.[7]

In line with the requirement that cognitive explanations need to capture the distinctions and inferences made by the cognitive system in question, Butterfill (2020), for instance, carefully develops an account of how infants acquire knowledge about simple facts that does not already presuppose that infants perceive facts. Butterfill clearly distinguishes between knowledge about facts and the underlying perceptual mechanisms, especially object indexing mechanisms. These perceptual object indexing mechanisms are not inferentially integrated with knowledge of simple facts. Such knowledge is acquired only later: "[O]bject indexes are independent of beliefs and knowledge states. Having an object index pointing to a location is not the same thing as believing that an object is there. And nor is having an object index pointing to a series of locations over time the same thing as believing or knowing that these locations are points on the path of a single object" (Butterfill, 2020, 66). Thus, Butterfill (2020) attempts to describe infants' cognitive capacities regarding facts in a way that captures the distinctions made within infants' cognitive apparatuses and that is true to the regularity expectations of infants. Thereby, Butterfill avoids simply alluding to our ontology.

To our understanding, Butterfill's (2020) attempt to spell out how infants structure their environment goes in the right direction but does not yet go far enough as it still centrally alludes to objects. In his argumentation, Butterfill centrally relies on Pylyshyn's and Leslie's accounts of object indexing (Pylyshyn, 2001, 2009; Leslie et al., 1998). There, object indexing is intended to explain how a sensory input that is not yet structured into objects is processed in a way that enables object individuation.

Butterfill (2020, 58) compares object indexing with assigning a pin on a map to the different trucks of a truck company in order to keep track of their current position. Each pin represents a truck, and the map represents the area in which the truck company

---

[7] Note that the phrase "being struck in perception" alone (by whatever) likewise suggests a factive interpretation. It entails (i) that something is perceived and (ii) that what is perceived is really the case (a fact).

operates. A pin's position on the map— which provides the spatial frame of reference— then represents its truck's position in the area at a given moment. According to this picture, individual objects are tracked by having assigned a particular index, just as each truck gets assigned a particular pin.

Note that an explanation of object individuation has to provide individuation criteria for objects. In common understanding, objects are individuated by the spacetime region they occupy during their existence (also cf. Leslie et al., 1998, 11). Consequently, object indexing accounts appeal to spatiotemporal coordinates in order to fix an index. These accounts, however, presupposes (i) that indexes are already individuated as one and the same at a certain moment (like the pins on the map) and (ii) that objects are already individuated (the trucks) or, at least, that we have a system of individuation criteria for objects at our disposal (a spatial frame of reference). Presupposing objects that are already individuated is clearly circular. But presupposing a spatial frame of reference is just as problematic. When a spatial frame of reference is already given, the problem of object individuation reduces to the question where the objects are and which ones there are. It does not answer the question of how a sensory input that is not yet structured into objects is processed in a way that enables reference to objects at all, that is, how individuation criteria for objects are acquired in the first place.

12.   Conclusion

Perner et al. have recently claimed that young children who still lack an understanding of mental states explain an agent's intentional action by appealing to the worldly facts that speak in favour of acting in that way. Perner et al. assume that thinking about worldly facts is to be conceptually distinguished from thinking about subjective perspectives, i.e., mental states, and that the first precedes the second. To argue that this cannot be right, we have distinguished two kinds of reference to facts, namely, reference *de re* and reference *de dicto*. While attributions of reference *de re* make no commitment to how a system or organism internally structures its environment, attributions of reference *de dicto* attempt to capture a given system's (implicit) ontology. Only attributions of reference *de dicto* are explanatory for cognitive capacities.

From a common-sense perspective, it seems plausible that young children can refer to worldly facts *de dicto* before they can refer to mental states *de dicto*. Facts, one could say, are simply out there in the world while mental states are hidden in people's heads. Due to the plausibility of this view, it seems, the conceptual complexity of reference to facts in the world is systematically underestimated and, correspondingly, infants' ability to refer to facts is systematically overestimated. However, as was argued above, objectivity and subjectivity are interdefined notions. Unless children have an

understanding that other people have different doxastic and spatial perspectives, they cannot have an understanding of facts either.

Our line of argument might appear to be undermined by the mere observation that infants are well able to react appropriately to what is going on in their environments. However, there are countless ways to refer *de re* to facts involving mental states and facts in the world in a predictable way. From the observation that children behave in line with expectations generated in our ontology (*de re*), it cannot be concluded that they have the *same* ontology (*de dicto*) (Hirsch, 1997, Hildebrandt et al., 2020a, 2020b). That is, similar behaviour can be constructed based on entirely different ontologies. This fundamental difficulty of all interpretations of children's behaviour should be taken seriously in early developmental contexts, where one has to assume that the ontologies *themselves* are under development.

The main argument was based on the analysis of how linguistic reference is made to facts, providing us with insights into the cognitive complexities involved in reference to facts *de dicto*. Notably, understanding spatial perspectives is necessary for understanding the special substitution rules of spatial demonstratives. The perspectival shifts that are inherent to the usage rules of, for instance, "here" and "there" lead to the construction of a single referent for different terms as they are used from different positions in space. "Here" and "there" can refer to the same location when uttered from different positions and the same position must be demonstratively referred to with different expressions from different positions. Understanding doxastic perspectives, in turn, surfaces in the ability to understand that any assertion can be true or false and that, correspondingly, SoAs can obtain or not obtain. In order to understand that assertions can be true or false, one must be able to consider SoAs without already taking them to be facts. One must be able to consider possible SoAs.

In combination, understanding these two perspectives means that the truth-value of a basic fact expressing assertion must be considered from different spatial perspectives. This comprises understanding that sentences can be evaluated differently from different spatial perspectives. Thereby, understanding spatial and doxastic perspectives constitutes understanding mental perspectives. In closing, it was argued that someone's (or something's) ability to use spatial indexicals appropriately can serve as a sufficiency criterion for whether that person (or system) refers to facts *de dicto*.

Note that understanding these perspectives is part of grasping factuality. Subjects who have not grasped subjective perspectives have not grasped factuality and cannot refer to facts *de dicto*. All they might be able to do is to systematically discriminate aspects of their environments that we refer to as different facts. They can only refer to facts *de re*. This finding is problematic for Perner et al. because merely attributing *de*

re reference to facts to young children is not explanatory and would lead to a homonymous use of "understanding" in "understanding worldly facts" and "understanding subjective perspectives".

A possible objection to this result would be to hold that a cognitive system might acquire an ontology by being presented in perception with facts. Thereby, reference *de dicto* to facts would allegedly be possible without understanding perspectives. Against this, it was argued that perceiving facts still requires being able to distinguish facts from non-facts, that is, from states of affairs that are not the case. This, in turn, amounts to understanding perspectives. The contrast between facts and non-facts (between truth and falsity) is part of the meaning of "fact". It is impossible to perceive facts *de dicto* without understanding perspectives unless one changes the use of the term "fact". Such a change, however, would have to be carefully clarified. Moreover, precautions would have to be taken not to rely on inferences that are only backed by the common usage of one's terms but not by one's reglementation. Perner et al.'s use of "fact" is not thus reglemented as the role of facts as objective reasons in Perner et al.'s account depends on truth.

While we are sympathetic to the overall endeavour to find a non-mentalistic theory of action explanation for adults, we are less optimistic that this theory also provides an account of its development. One cannot just remove the perspectival part and be left with a viable account of how children that do not yet understand perspectives explain intentional actions in terms of facts. Explaining young children's understanding of intentional action requires describing their cognitive abilities without recourse to reference to facts. That is, their abilities must be described in a way that captures the distinctions and inferences made by young children themselves. Similarly, pre-ToM children's abilities to form expectations about others' intentional actions would have to be explained without recourse to propositional attitudes.

To be sure, such an explanation would have to be made in our language. Nevertheless, we may not make the mistake of accidentally overloading the description with conceptual distinctions we as human adults use to navigate our environments. The challenge is to find terms of our language that allow for a description of the 'inner workings' of young children while avoiding misleading conclusions that are based on *our* ontology. What would be required for Perner et al. 's account as a developmental story is (i) an account of how children structure their environment before they learn to think with facts and perspectives and (ii) an explanation of how understanding facts and perspectives emerges from these abilities.

Our conjecture would be that young children's abilities to understand intentional behaviour are based on abilities of feature discrimination and feature pattern

generalization. When allowing for combined features, feature discrimination is a powerful starting point for learning all sorts of dependencies and regularities in one's environment (Hildebrandt et al., 2020a, 2020b). Expectations about natural as well as conventional regularities can be based on it such that it suffices for the acquisition of social norms, including the use of symbols. Reference to facts *de dicto* is then acquired via learning to refer to particular objects *de dicto*.

References

Apperly, I. A., and Butterfill, S.A. (2009). Do humans have two systems to track beliefs and belief-like states? *Psychological Review* 116, 4, 953–70.

Beckermann, A. (1996). Is there a Problem about Intentionality? *Erkenntnis*, 45(1), 1–23.

Bermúdez, J. L. (2007). *Thinking without words*. Oxford: Oxford University Press.

Burge, T. (2010). *Origins of objectivity*. Oxford: Oxford University Press.

Burr, D. C., Turi, M., & Anobile, G. (2010). Subitizing but not estimation of numerosity requires attentional resources. Journal of Vision, 10(6), 20–20. https://doi.org/10.1167/10.6.20

Butterfill, S. A. (2020). *The Developing Mind. A Philosophical Introduction*. London: Taylor & Francis.

Carey, S. (2009). *The origin of concepts*. Oxford University Press.

Casteñeda, H.-N. (1966). 'He': A study in the logic of self-consciousness. *Ratio*, 8, 130–157.

Casteñeda, H.-N. (1968). On the logic of attributions of self-knowledge to others. *Journal of Philosophy*, 65, 439–456.

Churchland, P. M. (1979). *Scientific Realism and the Plasticity of Mind*. Cambridge: Cambridge University Press.

Cohen, L. B., Chaput, H. H., & Cashon, C. H. (2002). A constructivist model of infant cognition. Cognitive Development, 17, 1323–1343. https://doi.org/10.1016/S0885-2014(02)00124-7

Davidson, D. (1963). Actions, Reasons, and Causes. *Journal of Philosophy*, 60 (23), 685-700.

Davidson, D. (2001). What is Present to the Mind. In D. Davidson, *Subjective, Intersubjective, Objective*. Oxford University Press.

Dennett, Daniel C. (1982). Styles of mental representation. *Proceedings of the Aristotelian Society*, 83, 213–226.

Dennett, D. (1987). Beyond Belief. In D. Dennett, *The Intentional Stance*. Cambridge (MA): MIT Press.

Evans, G. (1982). *Varieties of reference*. Oxford: Oxford University Press.

Field, H. (1981). Mental Representation with Postscript. In H. Field, *Readings in the Philosophy of Psychology, Vol. II* (pp. 78–114). Cambridge (MA): Harvard University Press.

Garson, J. W. (2013). *Modal logic for philosophers (2. ed)*. New York (NY): Cambridge University Press.

Glock, H. J. (2000). Animals, Minds and Concepts. *Synthese,* 123, 35–64.

Haith, M. M. (1998). Who put the cog in infant cognition? Is rich interpretation too costly? Infant Behavior and Development, 21(2), 167–179. https://doi.org/10.1016/S0163-6383(98)90001-7

Hildebrandt, F., Glauer, R., & Kachel, G. (2020a). Coming from a world without objects. *Mind & Language*, 1–18. https://doi.org/10.1111/mila.12313

Hildebrandt, F., Lonnemann, J., & Glauer, R. (2020b). Why Not Just Features? Reconsidering Infants' Behavior in Individuation Tasks. *Frontiers in Psychology*, *11*, 564807. https://doi.org/10.3389/fpsyg.2020.564807

Hirsch, E. (1997). Basic objects: A reply to Xu. *Mind & Language*, 12(3–4), 406–412. http://dx.doi.org/10.1111/1468-0017.00054

Hume, D. (1978). *A treatise on human nature*. (LA. Selby-Bigge, Ed.) Oxford: Oxford University Press.

Kaplan, D. (1977), Demonstratives: An essay on the semantics, logic, metaphysics, and epistemology of demonstratives and other indexicals. In J. Almog, J. Perry, H. Weinstein (Eds.), *Themes from Kaplan*, 481–564, New York: Oxford University Press.

Krøjgaard, P., Kingo, O. S., & Staugaard, S. R. (2013). Object individuation or object movement as attractor? A replication of the wide-screen/narrow-screen study by means of (a) standard looking time methodology, and (b) eye tracking. Child Development Research, 2013, 1–14. https://doi.org/10.1155/2013/639702

Machery, E. (2009). *Doing without concepts*. Oxford University Press.

Matthews, R. J. (2007). *The Measure of Mind*. Oxford: Oxford University Press.

Nagel, T. (1974). What Is It Like to Be a Bat? *Philosophical Review*, 83, 435–450.

Nagel, T. (1986). *The view from nowhere*. New York: Oxford University Press.

Nelson, M. (2019). Propositional Attitude Reports. In E. N. Zalta (Ed.), *The Stanford Encyclopedia of Philosophy* (Spring 2019). Metaphysics Research Lab, Stanford University. https://plato.stanford.edu/archives/spr2019/entries/prop-attitude-reports/

Objectivity (n.d.). Oxford Learner's Dictionary: English. Retrieved from: https://www.oxfordlearnersdictionaries.com/definition/english/objectivity, 08.01.2020.

Oxford Dictionary of English (3rd ed.). (2010). Oxford University Press. https://doi.org/10.1093/acref/9780199571123.001.0001

Peacocke, C. (1995). *A study of concepts*. MIT Press.

Perner, J. (1987). Three-year-olds' difficulty with false belief: The case for a conceptual deficit. *British Journal of Developmental Psychology*, 5, 125–137.

Perner, J., & Roessler, J. (2010). Teleology and causal reasoning in children's theory of mind. In J. Aguilar, A. Buckareff (Eds.), *Causing human action: New perspectives on the causal theory of action* (pp. 199–228). Cambridge (MA): MIT Press.

Perner, J. & Roessler, J. (2012). From infants' to children's appreciation of belief. *Trends in Cognitive Sciences* 16(10), 519–525.

Perner, J. & Esken, F. (2015). Evolution of human cooperation in Homo Heidelbergensis: Teleology versus mentalism. *Developmental Review*, 38, 69–88.

Perner, J., Priewasser, B., & Johannes Roessler, J. (2018). The practical other: teleology and its development. *Interdisciplinary Science Reviews*, 43(2), 99–114.

Perry, J. (2000). *The Problem of the Essential Indexical and Other Essays, Expanded Edition*, Stanford (CA): CSLI Publications.

Piaget, J. (1963). *The Origins of Intelligence in Children*. W.W. Norton.

Priewasser, B., Rafetseder, E., Gargitter, C. & Perner, J. (2018). Helping as an early indicator of a theory of mind: Mentalism or Teleology? *Cognitive Development*, 46, 69–78.

Pylyshyn, Z. (2001). Visual indexes, preconceptual objects, and situated vision. *Cognition*, 80(1), 127–158.

Pylyshyn, Z. (2009). *Things and Places: How the Mind Connects with the World*. Cambridge, MA: MIT Press.

Quine, W.V.O. (1960) *Word and object*. Cambridge (MA): MIT Press.

Reiss, J., Sprenger, J. (2017). "Scientific Objectivity", In E. Zalta (Ed.), *The Stanford Encyclopedia of Philosophy* (Winter 2017 Edition). https://plato.stanford.edu/archives/win2017/entries/scientific-objectivity/

Roessler, J. & Perner, J. (2013). Teleology: Belief as perspective. In S. Baron-Cohen, H. Tager-Flusberg, & M. Lombardo (Eds.), *understanding other minds*, 35–50, Oxford: Oxford University Press.

Russell, B. (1905). *On denoting*. Mind, 14(56), 479–493. http://www.jstor.org/stable/2248381

Smith, L. B., & Thelen, E. (2003). Development as a dynamic system. Trends in Cognitive Sciences, 7(8), 343–348. https://doi.org/10.1016/S1364-6613(03)00156-6

Stavans, M., Lin, Y., Wu, D., & Baillargeon, R. (2019). Catastrophic individuation failures in infancy: A new model and predictions. Psychological Review, 126(2), 196–225. https://doi.org/10.1037/rev0000136

Stern, D. N. (1985). *The Interpersonal World of the Infant: A View from Psychoanalysis and Developmental Psychology*. New York: Basic Books.

Strawson, P. F. (1959). *Individuals*. London: Methuen.

Tarski, 1935, "Der Wahrheitsbegriff in den formalisierten Sprachen", *Studia Philosophica*, 1: 261–405. References are to the translation by J. H. Woodger as "The concept of truth in formalized languages" in Tarski (1983).

Tarski, 1944, "The semantic conception of truth", *Philosophy and Phenomenological Research*, 4: 341–375.

Tomasello, M. (2003). *Constructing a language: A Usage-Based Theory of Language Acquisition*. Cambridge, MA: Harvard University Press.

Tomasello, M. (2018). How children come to understand false beliefs: A shared intentionality account. *PNAS*, 115(34), 8491–8498. www.pnas.org/cgi/doi/10.1073/pnas.1804761115

Tomasello, M. (2019). *Becoming Human. A Theory of Ontogeny*. Cambridge (MA): The Belknap Press of Harvard University Press.

Trevarthen, C. (1979). Communication and cooperation in early infancy: A description of primary intersubjectivity. In M. Bullowa (Ed.), *Before Speech: The Beginning of Interpersonal Communication* (321–347). Cambridge: Cambridge University Press.

Tugendhat, E. (2016), *Traditional and analytical philosophy: Lectures on the philosophy of language*. Cambridge, MA: Cambridge University Press. https://doi.org/10.1017/CBO9781316535608

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

Wittgenstein, L. (1922). *Tractatus Logico-Philosophicus. Logisch Philosophische Abhandlung*. London: Kegan Paul. Side-by-side-by-side edition, version 0.58 (24.5.2020). http://people.umass.edu/klement/tlp/, last checked 10.11.2020.