The Complementarity of Mindshaping and Mindreading

Uwe Peters
Centre for Logic and Philosophy of Science, KU Leuven, Belgium
Department of Economics, University College London, UK

Abstract

Why do we engage in folk psychology, that is, why do we think about and ascribe propositional attitudes (PA) such as beliefs, desires, intentions etc. to people? On the standard view, folk psychology is primarily for *mindreading*, for detecting mental states and explaining and/or predicting people's behaviour in terms of them. In contrast, McGeer (1996, 2007, 2015), and Zawidzki (2008, 2013) maintain that folk psychology is not primarily for mindreading but for *mindshaping*, that is, for moulding people's behaviour and minds (e.g., via the imposition of social norms) so that coordination becomes easier. Mindreading is derived from and only as effective as it is because of mindshaping, not *vice versa*. I critically assess McGeer's, and Zawidzki's proposal and contend that three common motivations for the mindshaping view do not provide sufficient support for their particular version of it. I argue furthermore that their proposal underestimates the role that epistemic processing plays for mindshaping. And I provide reasons for favouring an alternative according to which, in social cognition involving ascriptions of propositional attitudes, neither mindshaping nor mindreading is primary but both are complementary in that effective mindshaping depends as much on mindreading as effective mindreading depends on mindshaping.

Introduction

We often think about and ascribe propositional attitudes (PAs) such as, for instance, beliefs, desires, and intentions to others and ourselves. For example, when we hear Jane assert 'It is raining', or when we see her reach for a glass of wine, we tend to ascribe to her the belief that it is raining, and the desire to drink some wine, respectively. When we ascribe PAs to people in these ways, that is, when we think (as opposed to merely say) that a subject *S* holds a particular PA, this is often called *folk psychology*.

The term 'folk psychology' is sometimes used to refer to our practice of ascribing mental states and viewing people as minded, in general. I here want to focus on what Hutto (2008) and Zawidzki (2008: 194) take to be the "stricto sensu", narrow meaning of the term, and this is the practice of ascribing PAs, in particular.[1]

In recent years, there has been an intriguing debate on the purpose of folk psychology. Two different views on the matter can be distinguished, the *mindreading view*, and the *mindshaping view*.

The mindreading view holds that folk psychology serves mindreading. That is, it has the function of allowing us to determine people's pre-existing PAs and to explain and/or predict people's behaviour in terms of them. This might involve theorising about people's minds, simulating their perspective, or both (Baron-Cohen, 1995: 12; Nichols and Stich, 2003: 3, 65f; Goldman, 2006: 3, 38f, 147f, 150; Carruthers, 2011).

The mindshaping view has a different focus. It holds that folk psychology has the function of enabling us to actively mould behaviour and minds so that people become more predictable, and coordination and cooperation with them become easier (Mameli, 2001; McGeer, 2007, 2015; Zawidzki, 2008, 2013).

The mindreading view and the mindshaping view are not mutually exclusive. In fact, advocates of the mindreading view typically grant that folk psychology also involves at least sometimes mindshaping. Similarly, advocates of the mindshaping view typically grant that folk psychology also involves at least sometimes mindreading (McGeer, 2007: 138f; Zawidzki, 2008: 194).

The difference between the two views is a "difference in emphasis" (Zawidzki, forthcoming: 3). More specifically, according to the standard version of the mindreading view, the *primary* function of folk psychology is *epistemic*; it is thought to serve mindreading. In contrast, according to the most developed version of the mindshaping view, which is the proposal defended by Victoria McGeer (1996, 2007, 2015), and Tadeusz Zawidzki (2008, 2013), the *primary* function of folk psychology is *practical*; it is thought to first and foremost serve mindshaping. Furthermore, mindreading is taken to be derived from and only as effective as it is because of mindshaping, not *vice versa* (McGeer, 2007: 139; 2015: 260; Zawidzki, forthcoming: 3).

---

[1] Folk psychology is also often understood as the "capacity for mentalistic attribution and explanation" (McGeer, 2015: 259), that is, it is taken to be an epistemic ability to "make sense" of people's behaviour by appealing to their PAs (Hutto, 2008: 3; Zawidzki, 2008: 194). However, this is not unproblematic when the function of folk psychology is at issue, because it already defines folk psychology so that it has the epistemic function of explaining behavior. We need a neutral term. Construing folk psychology only as the practice of ascribing PAs, as I do it here, avoids the problem of prejudging the matter.

McGeer's, and Zawidzki's proposal,[2] which has become popular in recent work on mental-state ascriptions (Strijbos and De Bruin, 2015; Huebner, 2016; Bohm, 2015; De Bruin, 2016), "reverses the priority" of the standard conception of folk psychology, and calls for a significant revision of this conception (Zawidzki, forthcoming: 3). Since that is so, it deserves careful attention and scrutiny. Yet, so far no critical discussion of the proposal exists in the literature.

In the following, I want to change this and critically assess McGeer's, and Zawidzki's account. I shall argue that three often-mentioned motivations for the mindshaping view do not provide sufficient support for their particular version of it. Furthermore, I contend that their proposal underestimates the role that epistemic processing plays for the efficacy of mindshaping, and I defend an alternative account that occupies a space in between the standard mindreading view and McGeer's, and Zawidzki's mindshaping view. It holds that in social cognition involving PA ascriptions neither mindshaping nor mindreading is primary but both are complementary in that effective mindshaping depends as much on mindreading as effective mindreading depends on mindshaping.

Section 1 specifies further what I shall mean by 'mindshaping', mentions three key motivations for the mindshaping view, and elaborates on McGeer's, and Zawidzki's particular proposal. Section 2 returns to the motivations for the mindshaping view and argues that they do not suffice for McGeer's, and Zawidzki's account. Sections 3 and 4 provide positive grounds to doubt their proposal and motivate the view that mindshaping and mindreading are complementary. Section 5 briefly summarises and concludes the discussion.

1. The mindshaping view

The term 'mindshaping' might be understood in many different ways. Zawidzki (2013: 29f) proposes the following general definition: mindshaping involves a relation between a target mind (i.e., the mind being shaped), a cognitive mechanism M that was selected for the function to shape minds, and a model that M works to make the target mind match. Mindshaping happens when M performs the function that it was selected for.

Understood in this general way, mindshaping occurs even in non-human animals, for example, in baby rats when they are learning which foods to favor based on odors that they detect on their mother's mouth (Galef et al., 1983; Zawdizki, forthcoming: 9). In the following, I shall focus only on human-specific mindshaping that takes place when we make people's behaviour and mental states comprehensible to us by modifying them so that they match more closely with our normative expectations on how to act and be minded (McGeer, 2007: 147, 2015: 260; Zawdizki, 2013: xiv).

This modification can happen in many different ways *via* practices and mechanisms such as pedagogy, norm cognition and enforcement, and self- and group-constituting narratives (McGeer, 2007; Hutto, 2008; Gallagher, 2012; Zawidzki, 2008, 2013). Furthermore, it might

---

[2] Notice that McGeer (2007, 2015) and Zawidzki (2008, 2013) do not defend entirely the same view. They both independently developed different proposals (see below, section 1.2). When I refer to their views here and below by using the singular, i.e., 'proposal', 'account' etc., I merely mean to highlight the fact that they both advocate a version of the mindshaping view that is characterized by the two features just mentioned. This is compatible with their accounts also being distinct in various other ways.

apply to different kinds of behaviour and mental states including non-propositional ones, and it need not involve PA ascriptions.

Nonetheless, current advocates of the mindshaping view are explicit that their proposal also pertains to folk psychology *stricto sensu* and so to PA ascriptions. They hold that PA ascriptions too serve to mould the ascribee's behaviour and mind so that intersubjective predictions of actions and coordination are facilitated (Mameli, 2001; McGeer, 2007, 2015; Zawidzki, 2008, 2013). To keep the discussion focussed, I shall henceforth take the terms 'mindshaping' and 'mindshaping view' to refer specifically only to the subclass of instances of mindshaping that pertain to PAs and involve PA ascriptions.

1.1 Motivations for the mindshaping view

Different points have been mentioned to motivate the mindshaping view. The following three are recurring and central ones.

The first one starts with the assumption that any finite set of behavioural evidence is in principle compatible with an infinite number of distinct PAs. For instance, my raising my hand might be indicative of an intention to greet you, of an intention to hail a taxi, to call a waiter, and so on. Conversely, an ascription of any finite set of PAs is in principle compatible with an infinite number of distinct behavioural predictions: no particular set of PAs entails any particular behavioural response (Zawidzki, 2008: 196; Morton, 1996). The argument is then that if folk psychology were merely an epistemic device to determine people's PAs, and explain and/or predict behaviour, then in real-world situations, a computational problem would arise for the production of swift and reliable PA ascriptions. This is because the relationship between observable behaviour and the PAs that presumably cause it would become too complex to enable timely, accurate predictions, as the search space of possible interpretations would be too extensive. Advocates of the mindshaping view hold that in order to put constraints on the search space and to make swift and reliable PA ascriptions *via* mindreading possible in the first place, mindshaping is needed. The thought is that socially instituted norms on how one is to act if one is taken to hold a certain PA prompt people to act and think in line with our expectations of and PA ascriptions to them, and this "drastically reduces the space of possibilities that we must consider in order to coordinate with our fellows" (Zawidzki, 2008: 202). Hence, "[m]oulding" behaviour and minds via PA ascriptions "helps solve a coordination problem that is otherwise intractable" (ibid: 199; McGeer, 2007: 149f).

How exactly do PA ascriptions help us make people act in ways we expect them to? This question brings me to a second, closely related and equally common point to support the mindshaping view. Suppose that PA ascriptions served only the determination of pre-existing PAs and the explanation and/or prediction of behaviour. If that were so, then in situations in which we find that an ascribee's behaviour is subsequently incompatible with our ascription of a particular PA to her, we should readily revise our ascription and replace it with an updated one. But, advocates of the mindshaping view continue, this is often not the case. When the ascribee acts out of line with our PA ascription to her, we often criticise her behaviour as somehow wrong rather than revise our ascription. For instance, when we ascribe a belief that *p* to *S* on the basis of her asserting that *p*, and soon afterwards find her assert that not-*p*, then we tend to sanction her behaviour and want her to change her action and thinking (McGeer, 2007: 148, 2015: 266, Zawidzki, 2008: 201f, 2013: 98).

Advocates of the mindshaping view claim that such "sanctioning makes most sense on the assumption that propositional attitude ascription functions to *shape* minds rather than *read* them" (Zawidzki, 2008: 201). The sanctioning is taken to indicate that when we use PA ascriptions, in our "everyday encounters, explanation and prediction are generally not our chief concern" (McGeer, 2015: 266).

A third point that proponents of the mindshaping view tend to mention to motivate their proposal is related to empirical findings suggesting that an initially inaccurate trait or PA ascription can become a *self-fulfilling prophecy*, that is, it can elicit social expectancies that then in turn cause the emergence of the traits and mental states ascribed. Here are three examples.

(1) Tom false believes that Jane is angry with him. Because of this belief, Tom behaves in a hostile way with Jane. This hostility makes Jane angry with Tom.

(2) Jack, a devout Christian, is convinced (but mistaken) that his young daughter, Sue, believes in God's existence. As a result, he acts as if she holds this belief, expects her to act accordingly (e.g., claim that God exists, read the bible, etc.), and is pleased if she does. Sue notices this and wishes to satisfy her father's expectations. After acting as if she believes in God for some time, Sue gradually comes to adopt the belief that God exists.[3]

(3) Will, a passionate cabinetmaker, wishfully believes that his 16-year old son, Ron, wants to be a cabinetmaker too. Will's belief about what his son wants to be leads him to act toward Ron with the expectation that he wants to be a cabinetmaker. He, for instance, gives his son a hand plane, chisel and other woodworking tools for his birthday(s), shows him how to use them, praises him for his effort in applying these tools, and so on. Ron enjoys the increased attention from his father, and to maintain it, devotes much time to practicing his skill at woodworking. This in turn gradually leads him to want to become a cabinetmaker himself.

The phenomenon captured in (1)-(3) has been empirically well documented in various situations involving trait and PA ascriptions to others (Mameli, 2001: 609f; Zawidzki, 2008: 199f, 2013: 229). More generally, a recent and extensive meta-analysis of studies pertaining to self-fulfilling prophecies has found that while they are often only moderate, delicate, and transitory, "self-fulfilling prophecies in person perception are real, reliable, and occasionally quite powerful" (Jussim 2017: 1, 8).

They also occur when self-ascriptions are involved. Studies found, for instance, that subjects who had (unknowingly) falsely ascribed decisions and desires to themselves afterwards acted in line with these PAs in an endeavour to actively work to confirm their self-ascriptions (Zawidzki, 2013: 231, 2016; Strijbos and De Bruin, 2015: 299f). Arguably, when subjects come to the view that they hold a certain PA, they therewith regard themselves as committed to act in appropriate ways and the desire to honor that commitment and maintain consistency in their behavior then inclines them to govern their actions in ways that may lead them to acquire the PA. This is in line with and offers further motivation for the view that PA ascriptions involve mindshaping.

---

[3] That we can acquire a belief that *p* via acting as if we hold the belief (e.g., via repeatedly entertaining and asserting that *p*) is in line with a Spinozan account of belief formation; for empirical evidence supporting the account and a critique, see Peters (2017).

It is, however, important to note that the mindshaping view comes in different flavours and strengths. Moving from the strongest proposal to weaker versions, one might, for instance, hold that

(1) PA ascriptions *only* involve and are for mindshaping, not mindreading.
(2) PA ascriptions involve and are *primarily* for mindshaping, but sometimes also involve mindreading.
(3) PA ascriptions involve and are for *both* mindreading and mindshaping equally.
(4) PA ascriptions involve and are primarily for *mindreading*, but sometimes involve mindshaping.

Do the three just mentioned motivations for the mindshaping view provide sufficient support for McGeer's, and Zawidzki's particular version of it? Before addressing this question, I need to say more on the details of their proposal. But to anticipate, their view is a version of (2), and the alternative account that I shall develop below is a version of (3).

1.2 Mindshaping First – The Primacy Thesis

Neither McGeer nor Zawidzki deny that folk psychology sometimes involve mindreading (McGeer 2007: 138f, 2015: 260; Zawidzki, forthcoming: 3). Their goal is rather to offer a corrective to the standard mindreading view according to which PA ascriptions serve *primarily* epistemic purposes such as mindreading. This view overlooks the significance of the constructive, practical aspect of PA ascriptions, or so McGeer and Zawidzki hold.

For instance, McGeer (2007) writes that "very often when we make attributions" of "psychological states", including PAs, to "one another or to ourselves, we are not engaged in the activity of explaining and/or predicting behaviour [i.e., in an epistemic process] at all. We are engaged in the activity of moulding behaviour" (149). Despite playing an "epistemic" role too, in forming PA ascriptions, "explanation and prediction generally remain a *secondary* concern"; our "chief concern" is to shape behaviour and minds, McGeer (2015: 266) writes. She "rejects" the view that mindshaping is "built on the capacity to attribute mental states in the first place to try to figure out what others are up to – i.e., to try and explain and predict their behaviour" – and instead changes the ordering (McGeer, 2007: 139). Folk-psychology does not "involve a primary capacity for discovering or detecting (pre-existing) mental states", but a "primary capacity for forming and regulating our mental states" (McGeer, 2015: 260). The idea is that the epistemic capacity to determine pre-existing PAs and make sense of people's behaviour in terms of them is built on and "secondary" to the practical capacity to shape people's minds via PA ascriptions, not *vice versa* (McGeer, 2015: 266).

In the same vein, Zawidzki (forthcoming) writes that while one might hold that "[w]ithout a capacity to correctly represent independently constituted beliefs and desires we could not shape each other's minds as effectively as we do, […] the mindshaping hypothesis reverses this priority" (3). Zawidzki, just as McGeer, thus rejects the view that PA-specific mindreading and mindshaping are complementary in that they both depend on each other for their effectiveness. He holds that PA ascriptions that involve mindreading are "derived" from those used for mindshaping (2008: 205), and that the "reason why" the ability to produce PA ascriptions "was selected for and persists is mindshaping rather than mindreading" (2008: 194).

Notice that a trait might have been selected for and so have the evolutionary function of playing a particular causal role even if it does not typically play that role in everyday life. That is, the

question about the evolutionary function of a trait is distinct from the question of the frequency of its employment in everyday life.

Even though advocates of the mindshaping view have so far not explicitly noted it, the distinction between evolutionary function and frequency matters for the discussion of the mindshaping view. For one might hold that PA ascriptions were selected for (and so have the evolutionary function of serving) mindshaping while granting that in everyday life, PA ascriptions do not typically play that role, and in fact always involve and require epistemic processing such as the tracking of pre-existing PAs.

While this is a possible proposal, what matters here is only that it is not the view that McGeer and Zawidzki endorse. For, as noted, McGeer (2007) writes explicitly that in everyday life "attributions of a range of psychological states", including PAs, "*very often*" do not involve epistemic processing to track pre-existing "psychological states" and explain "behaviour" in terms of them "*at all* [emphasis added]" (149). That is, McGeer maintains that PA ascriptions very frequently do not involve the tracking of mental states, in general, and PAs, in particular, but typically involve mindshaping.

Zawidzki (2008, 2013) is less committal on the frequency of mindshaping (vs. mindreading) in PA ascriptions and more focused on the evolutionary function of them. But he too still writes that "[o]nce the use of mental state ascription to mind shape is reliable and prevalent, a derivative mind-reading use" of these ascriptions "is possible" (Zawidzki, 2008: 205). This suggests that on his view, PA ascriptions do not require any epistemic, mindreading processing pertaining to PAs. Rather, the latter is built on 'reliable and prevalent' mindshaping processing.

That is, in a nutshell, McGeer and Zawidzki commit to the following *primacy thesis* (PT):

> (PT) PA ascriptions involve primarily a non-epistemic, practical capacity for mindshaping, not an epistemic capacity for tracking pre-existing PAs (e.g., the mindreading faculty), and any epistemic capacity for tracking pre-existing PAs is derived from and only as effective as it is because of (PA-specific) mindshaping, not *vice versa*.[4]

In the remainder, I want to cast some doubts on (PT). I shall first consider whether the abovementioned three motivations for the mindshaping view provide sufficient support for it.

2. Revisiting the motivations for the mindshaping view

The first of the three motivations was that since any finite set of behavioral evidence is in principle compatible with an infinite number of distinct PAs, and an ascription of any finite set of PAs is in principle compatible with an infinite number of distinct behavioural predictions, the search space of possible interpretations of a particular behaviour would be too extensive to quickly arrive at reliable PA ascriptions, if the formation of PA ascriptions only involved epistemic processing. The claim then was that mindshaping is needed to deal with this computational problem and to enable mindreading, because it "drastically reduces the space of possibilities that we must consider in order to coordinate with" and make sense of our fellows;

---

[4] Notice that (PT) is compatible with the view that PA-specific mindshaping relies on various kinds of epistemic processing that does not involve the tracking of PAs but does involve the tracking of some other mental state and/or behavioural patterns. Indeed, this is arguably the view that Zawidzki (2013), in particular, favors. (PT) only pertains to the claim as to whether PA-specific mindshaping involves *PA-specific* epistemic processing, i.e., the tracking of PAs (as opposed to other mental states).

"[m]oulding" behaviour and minds thus "helps solve a coordination problem that is otherwise intractable" (Zawidzki, 2008: 202, 199; 2013; McGeer, 2007: 149f).

If successful, this point would lend some support to (PT). It is less convincing, however. I agree that practices of mindshaping, including the use of PA ascriptions, might play a role in solving the problem at issue. But it is arguably not the case that these practices offer the *only* solution to the problem, and that the latter would be 'otherwise intractable'. For instance, Andrews (2017) notes that in social interactions we often make sense of and coordinate our actions with others by relying (frequently unconsciously) on social stereotypes, which constrain possible interpretations of a target's behaviour and mental states by leading us to focus on some and ignore other information about the target.

One response here might be to propose that social stereotypes are reliable only in virtue of mindshaping mechanisms.[5] However, this proposal overlooks that stereotypes often track pre-existing facts of social reality (Gendler 2011). And by surveying the relevant empirical literature, Jussim (2012, 2017) has convincingly argued that the accuracy of many stereotypes is unlikely to be grounded in mindshaping processes such as self-fulfilling prophecies but due to people's sensitivity to pre-existing base rates. Since stereotypes reduce computational complexity and might facilitate reliable PA-involving social interpretations without involving mindshaping, it is not obvious that mindshaping is *required* to "reduce the space of possibilities that we must consider in order to coordinate with" and make sense of others (Zawidzki, 2008: 202, 199; McGeer, 2007: 149f).

It is worth noting too that, for instance, Gendler (2011) and Antony (2016) contend that implicit biases, which are largely unconscious and automatic evaluations that involve stereotype-based associations[6] between social groups and positive or negative properties (Brownstein and Saul, 2016),[7] might have similar epistemically beneficial effects in simplifying computational complexity in social environments and (sometimes) accurately tracking aspects of social reality. In fact, implicit biases are one kind of cognition that we might appeal to in order to motivate the view that largely unconscious, automatic processes may put into place a pre-established framework within which computations about a subject's PAs are constrained already from the beginning and, under normal conditions, reliable. This proposal does not rely on the assumption that the computations are limited by mindshaping, construed as a developmental process interfering with and changing the object of the computations. It rather rests on the assumption that when subjects approach the social world they always do so already equipped with a range of automatically operative parameters (e.g., implicit cognitions)[8] which are built into the architecture of the cognitive system and limit the scope of social inferences in a way that is sensitive to pre-existing facts.

Notice that the point here is not to argue that that proposal is correct. The point is rather that there are *prima facie* plausible suggestions on how the mentioned computational problem might be solved that do not require the postulation of mindshaping. Since stereotypes, biases, and other automatically operative cognitions might limit the search space for possible social

---

[5] I'm grateful to an anonymous reviewer here for pressing me on this point.

[6] There is debate on whether implicit biases are associations or beliefs (Brownstein, 2015). Nothing hinges on the matter here though.

[7] The empirical and philosophical research on implicit biases is not uncontroversial; see Peters (2018) for details and discussion.

[8] One such built-in cognitive feature to tackle computational intractability might be egocentric bias, see Peters (2016) for details and discussion. Another one might be the tendency to automatically accept the propositions one entertains, see Gilbert et al. (1993); but see also Peters (2017) for a critical discussion.

interpretations and incline subjects in PA-specific mindreading toward reliable conclusions even if there is no such a thing as mindshaping, it becomes questionable whether mindshaping is needed for solving "a coordination problem that is otherwise intractable" (Zawidzki, 2008: 202, 199; 2013; McGeer, 2007: 149f). Thus, the first point of motivation for the mindshaping view lends insufficient support to the proposal that mindreading is dependent on mindshaping. By extension, it does not sufficiently support (PT).

Consider now the second point. It was that often when we ascribe PAs to people and they subsequently act in a way incompatible with our ascriptions, we tend to criticise them for their action rather than simply revise our ascriptions. McGeer and Zawidzki hold that such "sanctioning makes most sense on the assumption that propositional attitude ascription functions to shape minds rather than read them" (Zawidzki, 2008: 201; McGeer, 2015: 266).

I agree that the regulative aspect of PA ascriptions suggests that PA ascriptions do not only involve mindreading but also mindshaping. However, the claim that the sanctioning at issue "makes most sense" on the assumption that PA ascriptions function to "shape minds rather than read them" is perhaps too strong (Zawidzki 2008: 201). The sanctioning makes equally good sense on the assumption that PA ascriptions serve *both* the shaping and reading of minds. In fact, the existence of the kind of sanctioning at issue might be taken to suggest that mindshaping is only operative *alongside* mindreading so as to ensure that the ascriber's initial mindreading-based explanations and predictions of a subject's behaviour are reliable. On this view, which is not ruled out by the consideration that McGeer and Zawidzki offer, the mindshaping effect of PA ascriptions sub-serves and facilitates the mindreading involved in the formation of PA ascriptions. Since this proposal suggests either a mindreading-mindshaping complementarity or a primacy of mindreading to mindshaping, it is at odds with (PT). The second point to motivate the mindshaping view thus also does not yield enough support for McGeer's, and Zawidzki's specific proposal.

The third and last point I mentioned pertained to self-fulfilling prophecies. The empirical data on the latter do make it plausible to assume that PA ascriptions do not always merely serve mindreading but also have constructive features. However, the data suggest at best that PA ascriptions, no matter whether they are accurate or not, can have self-fulfilling, mindshaping effects. This is compatible with the view that PA ascriptions also always involve an epistemic capacity to determine PAs and that mindreading and mindshaping are complementary, which is at odds with (PT).

The three common motivations for supporting the mindshaping view that I introduced hence do no provide enough support for McGeer's, and Zawidzki's particular version of it. In the next two sections, I shall offer positive reasons to doubt their proposal and to favour a complementarity of mindshaping and mindreading.

3. PA ascriptions and commitments to their truth

McGeer and Zawidzki hold that

> (PT) PA ascriptions (i) involve primarily a non-epistemic, practical capacity for mindshaping, not an epistemic capacity for tracking pre-existing PAs (e.g., the mindreading faculty), and (ii) any epistemic capacity for tracking pre-existing PAs is derived from and only as effective as it is because of (PA-specific) mindshaping, not *vice versa*.

In this section, I want to focus on (PT) (i). In the next section, I shall then address (PT) (ii).

It would contradict (PT) (i), if it turned out that PA ascriptions involved and depended upon an epistemic capacity for tracking pre-existing PAs. There is reason to believe that they do. Notice first that in general, in quotidian[9] uses of folk psychology, when we form a PA ascription, we affirm and commit to the truth of a proposition that attributes a PA to a subject *S* such as, for instance, the proposition *S believes that p*. That is, we do not only utter linguistic attributions, for instance, sentences of the form '*S* believes that *p*', which would not require us to commit to the truth of any proposition about *S*'s mind. Rather, we form judgments about people's minds, which do require some commitment to the truth of propositions about their minds.

McGeer and Zawidzki agree with this specification of quotidian folk psychology. For the examples that they mention when they introduce the term 'folk psychology' are cases in which the ascriber "concludes" that the ascribee is in a certain state of mind: they are cases in which the ascriber judges a proposition about an ascribee's mind to be true (McGeer, 2015: 259; Zawdizki, 2008: 193).

In any case, if one took the term 'folk psychology' to refer to linguistic expressions about people's PAs rather than judgements then one would no longer be talking about the same explanandum as advocates of the mindreading view do. This is because advocates of the mindreading view focus on committal thinking about PAs, not merely on our talking about them (see, e.g., Nichols and Stich, 2003; Carruthers, 2011). Hence, in order to stay focussed on the same explanandum in the debate with advocates of the mindreading view, proponents of the mindshaping view should agree that the phenomena whose function is to be explained are *committal thoughts* about PAs, that is, judgments about them.

In the debate at issue, advocates of the mindshaping view then hold that the primary function of these judgments is mindshaping whereas advocates of the mindreading view hold that their primary function is mindreading. Suppose that this much is agreed upon. Suppose too that PA ascriptions do involve a commitment to the truth of a proposition that attributes a PA to *S*.

Committing to the truth of a proposition *p* typically requires one to detect some evidence in support of *p*, because making such commitment and, more generally, forming a judgment is not a matter of direct choice. I cannot judge that *p* merely by deciding to judge that *p*. Direct doxastic voluntarism is arguably false (Booth, 2007: 115-130; Buckareff, 2014). If it is false, then since PA ascriptions involve a commitment to the truth of a proposition that ascribes a PA to *S*, there is ground to hold that PA ascriptions involve an epistemic capacity that tracks evidence of pre-existing PAs.

McGeer and Zawidzki might respond in the following three ways.

Response (1) They could grant that PA ascriptions involve a commitment to the truth of a proposition that attributes a PA to *S*, but then argue that this commitment need not be based on the ascriber's tracking evidence of pre-existing PAs, as opposed to behavioural or social facts. They might maintain that the commitment rests rather on evidence that the ascribee has acted in ways that confer on her certain PA-specific obligations that are captured by social norms of rationality whose grasp constitutes a form of "know-how" (McGeer, 2007, 2015). For instance,

---

[9] The term 'quotidian' is meant as a qualifier, because we often use folk psychology in fictional contexts, e.g., in story telling to describe a protagonist's action. In these cases, we do not commit to the truth of our PA ascriptions.

McGeer and Zawidzki hold that "full-blown propositional attitude attribution involves the undertaking and attribution of discursive commitments" (Zawidzki, 2013: 173; see also McGeer, 2007: 146). They could thus argue that when *S* claims that *p*, and we ascribe to her the belief that *p*, this ascription simply highlights that given what she said, *S* now has certain obligations and assertional commitments to the claim that *p*. *S*'s overt verbal behaviour provides the evidence for our taking it to be true that she believes that *p*, where this only means that we take it to be true that she has the mentioned obligations and commitments. If PA ascriptions are understood in this way, McGeer and Zawidzki could hold that while PA ascriptions do involve an ascriber's taking it to be true that an ascribee holds a certain PA, this does not require the capacity to track pre-existing mental states such as PAs. All that is needed is to monitor the ascribee's actions and to situate them within the context of the social norms regulating people's actions.

However, there is reason to believe that this proposal does not establish what it needs to. For suppose we hear *S* claim that *p* but at the same time also detect signs of her being unconvinced that *p*. Arguably, we would then refrain from ascribing to her assertorial commitments to the claim that *p* because we would not take her to be sufficiently certain about the truth of *p* and thus to be able to live up to the commitments and obligations that the belief ascription comes with. Our viewing *S* as having acted in ways that justify our assigning to her belief-specific commitments and obligations is hence not only dependent on the particular kind of action that *S* exhibits (e.g., her claiming that *p*) and the social norms that that action aligns with/contradicts, but also on our epistemically monitoring her for cues of the presence/lack of conviction, doubt etc. with respect to *p* (Sperber et al., 2010). That is, our viewing *S* as having the obligations and commitments at issue, and therewith our viewing her as holding the belief that *p*, arguably involves an epistemic capacity to track evidence of pre-existing PAs, namely her conviction, doubt, etc.

Response (2) In order to avoid the concession that PA ascriptions rely on such a capacity, McGeer and Zawidzki could argue that the ascriber's affirmation of the proposition that the ascribee holds a certain PA is based on the ascriber's *motivated reasoning*, which is the tendency to reason by accessing, constructing, and evaluating arguments/evidence in a biased manner so as to arrive at and adopt preferred conclusions (think of some people's denial of global warming, or their discrediting the science about it) (Kunda, 1990; Epley and Gilovich, 2016). Applied to the case at hand, the proposal might be that the ascriber's affirmation of the proposition that the ascribee holds a certain PA is based on the ascriber's desire that the ascribee act in certain ways. If that is right, it might seem that McGeer and Zawidzki can retain the proposal that even though PA ascriptions involve the ascriber's committing to the truth of a proposition about pre-existing PAs, this does not require an epistemic capacity to track evidence of PAs.

However, importantly, motivated reasoning is not evidentially unbound (Epley and Gilovich, 2016). Empirical research suggests that while people are indeed "more likely to arrive at conclusions that they want to arrive at," their "ability to do so is constrained by their ability to construct seemingly reasonable justifications for these conclusions" (Kunda, 1990: 480). Studies show that subjects only draw the

> desired conclusion if they can muster up the evidence necessary to support it. In other word, they maintain an 'illusion of objectivity'. To this end, they search memory for those beliefs and rules that could support their desired conclusion. They may also creatively combine accessed knowledge to construct new beliefs that could logically support the desired conclusion. It is this process of memory search and belief

construction that is biased by directional goals [i.e., goals unrelated to accuracy].

> The objectivity of this justification construction process is *illusory* because people do not realize that the process is biased by their goals, that they are accessing only a subset of their relevant knowledge, that they would probably access different beliefs and rules in the presence of different directional goals, and that they might even be capable of justifying opposite conclusions on different occasions. (Kunda, 1990: 483)

If this is right, and if we assume that mindshaping PA ascriptions are based on motivated reasoning then these ascriptions too are perhaps at least partly based on evidence pertaining to the presence of the PA ascribed. Otherwise the ascriber would have difficulties sustaining the 'illusion of objectivity' about the ascription. This is at odds with (PT), according to which PA ascriptions involve primarily a non-epistemic capacity for mindshaping, not an epistemic one for tracking PAs.

Response 3 McGeer and Zawidzki might hold that PA ascriptions have 'forward-looking' truth conditions capturing what will (as opposed to what already is) the case and involve the ascriber's being confident that her ascriptions will bring about the truth of the ascription. On this view, the ascriber can take her PA ascription to be true without having to use the capacity to track pre-existing mental states. McGeer (1996) and Zawidzki (2016) explicitly defend such a proposal when it comes to PA *self*-ascriptions. They hold that PA self-ascriptions have "forward-looking truth conditions, or 'satisfaction conditions'" and amount to knowledge because with them, we commit ourselves to regulate our own actions and thoughts so that they align with the ascriptions (McGeer, 1996: 508, 515). Zawidzki (2013: 231) adds that given that "third-person regulative interpretation is also possible", which is what happens, for instance, "when children accept and conform to the implications of their caretakers' interpretations of their behaviour", when we form PA *other*-ascriptions and commit to their truth, then this commitment too might be based on the regulative implications of our ascriptions, which we anticipate to become fulfilled.

However, notice that while the assumption of a 'forward-looking' insight into the truth of PA ascriptions may seem plausible in one's own case, it is much less plausible when it comes to PA other-ascriptions. This is because in one's own case, one has direct control over one's own action and thinking, yet when it comes to another subject, this is not case. As McGeer (1996) herself notes,

> I cannot make it the case that you behave in ways coherent with what I say you hope, desire, or fear any more than I can make it the case that the world is a certain way by announcing how (I think) it is; but I can and do govern my own actions in ways that fit with the claims I make about myself. (507)

Since that is so, the third way in which McGeer and Zawidzki might respond to the point that PA ascriptions to *S* involve a commitment to the truth of a proposition that attributes a PA to *S* is, just as the other two, not fully convincing.

The initial point remains. Assuming that (1) PA ascriptions typically involve a commitment to the truth of a proposition that ascribes a PA to a subject and that (2) this commitment needs to rest at least in part on evidence for the truth of that proposition, PA ascriptions arguably involve epistemic processing to track pre-existing PAs. This is at odds with (PT) (i), according

to which PA ascriptions involve primarily a non-epistemic, practical capacity for mindshaping and not an epistemic capacity for tracking pre-existing PAs.

I shall now argue that there is also ground to be sceptical about (PT) (ii), according to which any epistemic capacity involved in PA ascriptions is derived from and only as effective as it is because of (PA-specific) mindshaping, not *vice versa*.

4. Strong PA-specific mindshaping effects depend on mindreading

To argue for this claim, it will be useful to start with a concrete example. Suppose Jane wants her son, Tom, to be a gifted piano player and wishfully believes him to be just that. If Tom is aware that his mother just wants him to be a gifted piano player (for instance, if her actions reveal that she does not in fact believe that he is), then Jane's belief-related expectations about him will have a reduced impact on Tom. This is because he is likely to anticipate that his mother will be less disappointed if he does not conform to her wishful expectations than if she genuinely believed him to be a gifted piano player. For clearly, Jane's *wanting* Tom to be a gifted piano player, even if it is a strong desire, is still more compatible with him not fitting that description than her *belief* that he in fact does. As a result, Jane's trait ascription to Tom is less likely to create self-fulfillment and mindshaping effects. It will have stronger effects if it involves a firm belief on Jane's part about the boy, because this will for the child increase the normative force of the related expectations.

There is empirical support for this point. Psychological studies conducted in teacher/pupil settings suggest that it is especially teachers who are *certain* that a particular student is more intellectually capable than the rest of the class that elicit self-fulfilment effects from the student, not those whose views about the student are more flexible and held less confidently (Jussim, 1986). Relatedly, Swann and Ely (1984) found that when a perceiver formed an expectancy about a target that was discrepant with that target's self-conception, self-fulfilling prophecies tended to "occur only when perceivers were certain of their expectancies and targets were uncertain of their self-conceptions" (1287). These findings suggest that a trait or PA ascription is more likely to have strong mindshaping effects, if it has what I shall call *epistemic oomph*, that is, if the ascriber is convinced that the ascribee already holds the trait or PA ascribed.

That PA ascribers are likely to be more successful in shaping an ascribee's behaviour and mind if they hold an unwavering belief that their views on the ascribee are already true also coheres with data from studies that compared the effects of persuasion and attribution. For instance, in a well-known experiment, Miller et al. (1975) compared the littering behaviour of children who had repeatedly been told that they should be tidy (this was the *persuasion* condition) with that of children who had repeatedly been told that they were in fact tidy (this was the *attribution* condition). Interestingly, Miller et al. found that children who had been told that they were in fact tidy displayed a sharp and lasting decrease in their rate of littering, whereas children who had been told that they ought to be tidy showed only a moderate decrease in littering and then returned to littering at the same rate as the control group (see also Velleman, 2000: 358f; Mameli, 2001: 611).

With this in mind, there is reason to assume that if an ascribee, say, Tom, perceived a trait or PA ascription to him by an ascriber, say, Jane, merely as a prompt to the effect that he *ought to* act in certain ways so as to, for instance, satisfy Jane's desire, rather than as an expression of her belief about a pre-existing trait or PA, then the ascription would be less effective in shaping Tom's behaviour and mind. The PA ascription would more likely elicit the effects familiar from

the persuasion condition rather than that of the attribution condition in the study just mentioned. Similarly, if Jane took it that her trait or PA ascriptions are merely based on what she *wants* to be the case, from her own point of view too, the expectations about Tom that the ascription comes with would also only have the character of what Tom ought to do (to satisfy Jane's desire), not what he will do (given the trait or PA she ascribes to him). Since disguising this from Tom requires cognitive resources that Jane will often lack (von Hippel and Trivers, 2011), unless she genuinely believed in the trait or PA ascription, the latter's mindshaping effects are likely to be less pronounced.

If this is correct and strong mindshaping effects of Jane's trait and PA ascriptions depend on her taking the ascriptions to be true, then any mechanism that ensures that (1) Jane takes her ascriptions to Tom to track already existing traits or mental states and that (2) Tom detects her certainty about him will be important for the ascriptions to have strong mindshaping effects. The mindreading faculty, that is, the capacity to determine pre-existing PAs and explain and/or predict behaviour, is precisely such a mechanism. Its processing may not always only be sensitive to evidence relevant to an accurate representation of people's mind. It might often also be influenced by the ascriber's/ascribee's motivational states (Shepherd, 2012), values, or normative considerations, and the result might frequently be confabulations (Carruthers, 2011). But since its processing is still *epistemic* in nature – allowing Jane to take her PA ascriptions to be true and Tom to detect Jane's conviction – the mindreading faculty is likely to facilitate and be critical for strong mindshaping effects.

To be clear, mindshaping via PA ascriptions might well occur without the ascribers' taking their PA ascriptions to be true and without the ascribees' detecting the underlying degree of conviction. The point here is just that PA ascriptions are likely to have a *reduced* mindshaping impact if they do not involve the ascribers taking their PA ascriptions to be true and the ascribees tracking this. And for ascribers to take their PA ascriptions to be true and well motivated and for the ascribee to register this, the mindreading capacity will be important.

This point matters for the discussion of McGeer's, and Zawidzki's proposal, because they reject the view that PA-specific mindshaping is dependent on and plays an equally important role as PA-specific mindreading in social cognition. For them, PA ascriptions involve primarily mindshaping, and the latter facilitates mindreading, not *vice versa* (McGeer, 2007: 139; Zawdizki, forthcoming: 3). The preceding consideration challenges this view as it suggests that the effectiveness of PA ascriptions in mindshaping depends on the effectiveness of PA-specific mindreading. That is, while there is little doubt that mindshaping facilitates successful mindreading in that it helps make people act in line with our normative expectations (which in turn makes them more easily predictable and so readable for us), mindreading *in turn* facilitates effective mindshaping. This is because it provides PA ascriptions with the epistemic oomph that they need so as to not be interpreted (by ascriber and ascribee alike) in ways that would weaken the expectations they create.

5. Conclusion

McGeer and Zawidzki hold that PA ascriptions (i) involve primarily a non-epistemic, practical capacity for mindshaping, not an epistemic capacity for tracking pre-existing PAs (e.g., the mindreading faculty), and (ii) any epistemic capacity involved in PA ascriptions is derived from and only as effective as it is because of PA-specific mindshaping, not *vice versa*. My aim here was to scrutinise this view and motivate a revision of it by highlighting the involvement of epistemic processing in PA ascriptions that McGeer and Zawidzki seem to have overlooked.

I indicated doubts about (i) by arguing that since PA ascriptions typically involve a commitment to the truth of a proposition that attributes a certain PA to a subject, there is reason to believe that they involve an epistemic capacity for tracking pre-existing PAs. I then argued that, against the claim captured in (ii), there is also ground to hold that strong mindshaping effects depend on this epistemic capacity. For there is empirical evidence suggesting that PA ascriptions trigger strong mindshaping effects only if they are sincere and strongly held. Furthermore, there is reason to believe that this in turn presupposes a mindreading use of PA ascriptions. This is because, on the one hand, the ascriber needs to strongly believe that the ascribee already holds the PA attributed, and, on the other hand, the ascribee needs to detect this conviction in the ascriber. Both of these factors suggest that pronounced PA-based mindshaping effects rely on PA-based mindreading.

A position that is situated between both the standard mindreading view (which holds that PA ascriptions are primarily for mindreading), and McGeer's, and Zawidzki's mindshaping account (which holds that PA ascriptions are primarily for mindshaping) thus suggests itself. It holds that when it comes to the function of PA ascriptions neither mindreading nor mindshaping is primary or more important than the other. Rather, both are complementary and interdependent in that effective mindshaping depends on mindreading just as much as effective mindreading depends on mindshaping.

References

Andrews, K. (2017). More stereotypes, please! The limits of 'theory of mind' and the need for further studies on the complexity of real world social interactions. *Behavioral and Brain Sciences*, 40: 20–22.

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

Booth, A. (2007). Doxastic Voluntarism and Self-Deception. *Disputatio: International Journal of Philosophy* 2 (22): 115-130.

Brownstein, M. (2015). Implicit bias. *Stanford Encyclopedia of Philosophy*, Zalta, E. (ed.). <http://plato.stanford.edu/entries/implicit-bias/>.

Brownstein, M. & Saul, J. (2016). Introduction. In: Brownstein, M., and Saul, J. (eds.), *Implicit Bias and Philosophy, Volume 1: Metaphysics and Epistemology*. Oxford: OUP, pp. 1-20.

Buckareff, A. (2014). Deciding to Believe *Redux*, in *The Ethics of Belief: Individual and Social*, ed. Jonathan Matheson and Rico Vitz, Oxford: Oxford University Press: 33–50.

Carruthers, P. (2011). *The Opacity of Mind*. Oxford: Oxford University Press.

Galef, B. G., Wigmore, S. W., & Kennett, D. J. (1983). A failure to find socially mediated taste aversion learning in Norway rats (R. norvegicus). *Journal of Comparative Psychology*, 97(4), 358–363.

Gallagher, S. (2012) In defense of phenomenological approaches to social cognition: interacting with the critics. *Rev Philos Psychol* 3(2): 187–212.

Gendler, T. S. (2011). On the epistemic costs of implicit bias. *Philosophical Studies* 156: 33–63.

Gilbert, D., Tafarodi, R., and Malone, P. (1993). You can't not believe everything you read. *Journal of Personality and Social Psychology*, 65, 221–233.

Goldman, A. (2006). *Simulating minds*. Oxford: OUP.

Huebner, B. (2016). Implicit bias, reinforcement learning, and scaffolded moral cognition. *Implicit Bias & Philosophy: Volume I, Metaphysics and Epistemology*. M. Brownstein, & J. Saul (eds.). Oxford: Oxford University Press.

Hutto, D. (2008) *Folk psychological narratives: the sociocultural basis of understanding reasons*. MIT Press, Cambridge.

Jussim, L. (1986). Self-fulfilling Prophecies: A Theoretical Integrative Review. *Psychological Review*, 93: 429–445.

Jussim, L. (2012). *Social Perception and Social Reality*. Oxford: Oxford University Press.

Jussim, L. (2017). Précis of social perception and social reality: Why accuracy dominates bias and self-fulfilling prophecy. *Behavioral and Brain Sciences*, 40, 1–20.

Kunda, Z. (1990). The Case for Motivated Reasoning. *Psychological Bulletin*, 108, 3: 480-498.

Mameli, M. (2001). Mindreading, mindshaping, and evolution. *Biology and Philosophy*, 16: 597–628.

McGeer, V. (1996). Is 'self-knowledge' an empirical problem? Renegotiating the space of philosophical explanation. *Journal of Philosophy*, 93: 483–515.

McGeer, V. (2007). The regulative dimension of folk psychology. In: *Folk psychology re-assessed*, (eds.) Hutto, D. and Ratcliffe, M., Dordrecht: Springer: 138–56

McGeer, V. (2008) The moral development of first-person authority. *European Journal of Philosophy*, 16(1): 81–108

McGeer, V. (2015) Mind-making practices: the social infrastructure of self-knowing agency and responsibility. *Philosophical Explorations*, 18, 2: 259–281.

Nichols, S., and Stich, S. (2003). *Mindreading*. Oxford: OUP.

Peters, U. (2016). Human Thinking, Shared Intentionality, and Egocentric Biases. *Biology & Philosophy*, 30, 6, 1–16.

Peters, U. (2017). On the Automaticity and Ethics of Belief. *Ethics, Law, and Cognitive Science*, (eds.) Mario De Caro and Massimo Marraffa, 99–115. Available at: https://uwepeters.weebly.com/papers.html

Peters, U. (2018). Implicit bias, Ideological Bias, and Epistemic Risks in Philosophy. *Mind & Language*. 1–27, DOI: 10.1111/mila.12194

Shepherd, J. (2012). Action, Mindreading and Embodied Social Cognition. *Phenomenology and the Cognitive Sciences*, 11, 4: 507–518.

Strijbos, D., and de Bruin, L. (2015). Self-Interpretation as First-Person Mindshaping: Implications for Confabulation Research. *Ethical Theory and Moral Practice*, 18 (2): 297–307.

Swann, W., and Ely, R. (1984). A battle of wills: Self-verification versus behavioral confirmation. *Journal of Personality and Social Psychology*, 46: 1287–1302.

von Hippel, W., Trivers, R. (2011) The evolution and psychology of self-deception. *Behavioral and Brain Sciences*, 34(1): 1–16.

Zawidzki, T. (2008). The function of folk psychology: Mind reading or mind shaping? *Philosophical Explorations*, 11(3): 193–210.

Zawidzki, T. (2013). *Mindshaping: A New Framework for Understanding Human Social Cognition*. Cambridge, MA: MIT Press.

Zawidzki, T. (2016). Mindshaping and Self-Interpretation. In: *The Routledge Handbook of Philosophy of the Social Mind Routledge* (ed.) Kieverstein, J., NY: Routledge.

Zawidzki, T. (forthcoming). Mindshaping. In: *The Oxford Handbook of 4e Cognition* (eds.) Albert Newen, Leon De Bruin, Shaun Gallagher. Oxford: OUP. Preprint retrieved from https://cear.gsu.edu/files/2016/09/Zawidzki_Mindshaping-Chapter.pdf