

Normativity-based Autonomy of Folk Psychology Reconsidered

Taavi Laanpere

Department of Philosophy, University of Tartu

It has been a recurring theme in the philosophy of mind that folk psychology is autonomous. This paper has three goals. First, it aims to clarify what the term ‘folk psychology’ could mean in different contexts. Four widespread senses of the term are distinguished and the one eligible for autonomy is picked out. Secondly, a classic argument for autonomy is introduced and motivated. This is the argument from the normativity of folk psychology, based on its constitutive rationality. According to this argument, mentalistic concepts are to be understood as components of prescriptions for a rational course of action, rather than descriptions. Thirdly, limits of the argument from normativity are demonstrated. At best, the argument applies to merely a small segment of explanations in terms of mentalistic vocabulary, as the latter is meant to convey much more than simply normative content about the rational profile of an agent.

Keywords: folk psychology, autonomy, normativity, rationality

1. Introduction

A common theme in the philosophy of mind is that folk psychology is *special*—both in that it is in some important way distinct from the enterprise of scientific psychology and that its relation with scientific psychology is dramatically unlike, say, the relation between folk biology and scientific biology. At times, this amounts to an explicitly stated and defended view. In a philosophical tradition harking back to later Wittgenstein and Ryle, we can find elaborate arguments claiming that our everyday mentalistic vocabulary constitutes a self-sufficient discourse about human agents—a discourse which is in an important sense autonomous from, say, any neuropsychological description of those agents. Other times, such assumptions of distinctiveness

Corresponding author's address: Taavi Laanpere, Department of Philosophy, Institute of Philosophy and Semiotics, University of Tartu, Ülikooli 18, 50090 Tartu, Estonia. Email: taavi.laanpere@ut.ee.

are working more implicitly on the background. For instance, I suspect that they often underlie philosopher's armchair-practice of conceptual analysis, which is fueled by the notion that there is something interesting and important to be said about ordinary mental concepts, irrespective of how they hook up with scientific psychology (Jackson and Pettit 1990). While the conceptual repertoire of, say, folk biology is hardly ever found intrinsically interesting,¹ discussion of various familiar mental faculties has dominated much of contemporary philosophy.

One recurring line of reasoning behind this assumption of distinctiveness is the following.² Our folk-psychological vocabulary forms a certain kind of unity—a wholly distinct level of explanation, often called the 'personal level' (Bermúdez 2005, 34). This level of explanation is distinguished (possibly among other definitive features) by its inherent normativity. More specifically, the statements made at this level are often said to bear their normative character due to their connection with the necessary background assumptions about rationality, as explicated in the works of Donald Davidson (1974), Daniel Dennett (1987), John McDowell (1998) and others.³ Ascriptions of mental states—in stark contrast with ascriptions of physiological states—make sense only within the framework of *rational reconstruction* of an agent's behavior. In the folk-psychological mode of explanation, we ascribe states that the subject *rationally ought* to have in light of her behavior and other mental states. There is a radical discontinuity between this mode and other forms of explanation that are not sensitive to such normative principles—thus, folk psychology, being in a wholly different kind of explanatory business than e.g., scientific psychology, is autonomous from those other explanations. In turn, this explanatory autonomy is supposed to have repercussions for reductionist and eliminativist projects, and maybe even for physicalism in general (see e.g., Hornsby 1997).

¹ Which is not to say that it could not be of any (meta-)interest to a cognitive scientist studying the natural human reasoning about the organic world (see e.g., Medin and Atran 1999).

² Naturally, there are other cases to be made. For instance, one popular claim for autonomy is based on multiple realizability—i.e., the fact mental concepts seem to pick out abstract functional roles independently of their realization (see e.g., Jackson and Pettit 1990). Or, one might insist that ordinary mental concepts are special and autonomous, because they originate from introspection, i.e., they are “read off” one's own experience (see e.g., Goldman 2000). Here, I only intend to present a critique of the argument from normativity—these alternative cases for autonomy do not fall into the scope of this paper.

³ Again, this is not the only way to understand the normative nature of folk psychology. For example, according to one rather different explication of the normative element, folk psychology is essentially in the business of moral evaluation (Stueber 2009). Also, see (Mölder 2011, in Estonian) for an overview about various further ways in which the mental could be said to be normative. This paper focuses narrowly on the normativity from the alleged constitutive rationality.

In this paper, I will revisit this classic case for the autonomy of folk psychology. On the one hand, I want to maintain that there is something to its core idea. Contra general objections such as those presented by Jonathan Knowles (2002), I hold that an important *part* of the mentalistic discourse indeed stands out due to its normative character. On the other hand, however, I want to show that such arguments from normativity warrant a much weaker conclusion than they often are (at least implicitly) taken to. This is because our folk psychological vocabulary does not actually constitute a unity based on it being a system of *normative* distinctions—to the contrary, most of the distinctions between mental concepts are clearly due to their *descriptive* content. The take-home lesson of the paper is that we should be wary of using the normativity argument to conclude anything about the *whole* of folk psychology.

Before proceeding with this main task, however, some preliminaries regarding the very concept of ‘folk psychology’ need to be taken care of. This label has been used with a host of different meanings that have become notoriously entangled in the literature. Over the course of last decades, the dominant use has slowly shifted from a rather specific conceptual framework to a whole range of sociocognitive capacities. There is an emerging consensus that theorizing with mental concepts does not play the central role in social cognition it once was thought to play (see e.g., Ratcliffe and Hutto 2007; and other essays in that volume). Now, if we use the term ‘folk psychology’ in the latter sense (i.e., completely divorced from our mentalistic conceptual repertoire), then the autonomy claim is either nonsensical or trivially true, depending on how autonomy is understood. Thus, in section 2, I will start by presenting a taxonomy of meanings for ‘folk psychology’ in order to help with navigating these murky waters, and to elucidate (and legitimize) the sense of ‘folk psychology’ which is eligible for an interesting sort of autonomy in the first place. In section 3, I will introduce the idea of autonomy and the argument from the alleged normativity of folk psychology. I will also show that some principled attacks against this argument can be resisted. In section 4, however, I will switch sides and show that the argument from normativity nevertheless applies only to a very narrow range of folk psychological states. Consequently, any anti-reductionist or anti-eliminativist arguments drawn from such a narrow autonomy are likewise severely limited.

2. A road map for ‘folk psychology’

Let us start with two uncontroversial observations. First, people often think and speak of themselves and others in a shared vocabulary of various mental

states.⁴ Second, ordinary folk are normally quite adept in intersubjective understanding, coordinated group action, predicting each others' behavior, etc.,—briefly put, they have considerable sociocognitive capacities.

What makes the label 'folk psychology' often so frustrating is that based on merely those two observations, we can already state four different, although equally widespread meanings for it. On the one hand, the term could be (and has been) used to mean simply either of the observed phenomena. Thus, we could say that folk psychology is simply the "ordinary person's repertoire of mental concepts" (Goldman 2000, 15). Let us call this FP₁. Or, at the other end of a certain spectrum, we could save the label for the totality of our natural sociocognitive capacities—our "everyday, rather than 'scientific,' appreciation of mindedness" (Ratcliffe and Hutto 2007, 1). Let us call this FP₄.

Alternatively, 'folk psychology' might designate not the observed phenomena themselves, but rather the explanantia for either. For example, here is Peter Godfrey-Smith (2005, 1, my emphasis): "I take folk psychology to be the *basis*—whatever it is—of our ability to describe, interpret, and predict each other by attributing beliefs, desires, hopes, feelings, and other familiar mental states." Now, what is the basis of FP₁—where do our common mental concepts come from and what determines their meaning and reference? Probably the most popular answer of late 20th century to the latter question derives from the works of Wilfrid Sellars and David Lewis. Against the behaviorist hegemony of his time, Sellars (1956) constructed an extended thought experiment—the "Myth of Jones"—to show how and why one might legitimately come to use mentalistic vocabulary for inner, unobservable states. The key is to understand mental states as theoretical posits, which are introduced to explain the complex regularities in observable behavior. Such theoretical posits abound in science—think of subatomic particles and physical forces, for instance. Instead of being directly observable, their existence is inferred from systematically observed effects and a background theory used to explain these effects. Meanings of concepts like 'electron' and 'gravity' are constituted by the causal roles these posits play in the theories the concepts embedded in. According to a popular account, mental

⁴ Two comments—about the terms 'mental' and 'state,' respectively. First, there is no consensus over what constitutes the so-called mark of the mental. Popular candidates are intentionality and phenomenality. That is, something is mental if it is *about* something else and/or accompanied or constituted by a subjective feeling. I am not going to subscribe to any definite mark of the mental—by 'mentalistic' vocabulary and concepts, I simply mean those that are commonly held to designate mental phenomena. Second, I use the term 'mental *state*' here and throughout for stylistical convenience rather than as a substantial ontological description. I am not assuming that the folk are committed to existence of mental *states* as opposed to, say, mental processes, activities, properties or events.

terms likewise get their meanings and reference via their roles in the theory they are embedded in.

But what kind of theory are our common mental terms embedded in? Here's Lewis' proposal:

Think of common-sense psychology as a term-introducing scientific theory, though one invented long before there was any such institution as professional science. Collect all the platitudes you can think of regarding the causal relations of mental states, sensory stimuli, and motor responses. [...] Add also all the platitudes to the effect that one mental state falls under another—'toothache is a kind of pain,' and the like. Perhaps there are platitudes of other forms as well. Include only platitudes which are common knowledge among us—everyone knows them, everyone knows that everyone else knows them, and so on. For the meanings of our words are common knowledge, and I am going to claim that names of mental states derive their meaning from these platitudes. (Lewis 1972, 256)

This amounts to a position that has been variously called commonsense functionalism, analytical functionalism and the platitude view. While in the quote above Lewis invites us to think of such a term-introducing folk psychology as a theory "invented long before science," he does not mean to say that there actually was someone in the past who came up with a finite list of such causal principles, now regarded as platitudes. Like Sellars (1956, 48) with his Myth or "anthropological science fiction," he is making a point about how to plausibly interpret our present practice, rather than what happened in the actual history. Moreover, Lewis does not make any claims about how, to what extent or whether at all should such a theory be represented in the minds of each or any individual. Rather, it is an idealized construction, which emerges from our communal practice with the mentalistic vocabulary. Let us call this idealized collection of platitudes about causal connections between inputs, mental states and outputs FP_2 . On the Lewisian account, then, mental concepts of FP_1 derive their meanings and reference from being embedded in FP_2 . The term 'belief,' for instance, names the realizer of various roles for beliefs as specified by FP_2 and refers to the actual unique realizer of such roles, presuming that there is one (Lewis 1972).

Now, certain psychologists—commonly known as *theory-theorists*—have built on this idea and their own empirical studies to explain how mental state ascription (also known as "mindreading") and human social cognition in general actually work (see e.g., Gopnik and Meltzoff 1997). According to this once-popular account, all healthy adults share a tacit theory-like representation that underpins our ability to attribute mental states and thus successfully navigate in the social world. There is a range of different proposals regarding its exact form and process of acquisition, but the key idea is that

a theory roughly like the one Lewis describes is not merely a philosophers' abstraction, but a real body of information possessed by all competent mind-readers. Let us call this theory-like representation FP_3 .⁵

One should note that these four senses of folk psychology, while interconnected in various ways, can nevertheless come apart in some respects. Again— FP_1 is a set of concepts; FP_2 is part of a semantic hypothesis; FP_3 is part of an empirical hypothesis and FP_4 is simply a broad set of abilities. Most importantly for present purposes, the *fact* of us using FP_1 is independent of whether the hypotheses of FP_2 or FP_3 have any merit. This is important, because both hypotheses face well-known challenges. The FP_2 approach has been criticized for its commitment to descriptivism and its holistic implications. First, while Lewis suggested that the platitudes of FP_2 provide the definite descriptions that determine the reference of mental terms, such descriptivism regarding the reference of theoretical terms has been later powerfully argued against in Kripke's and Putnam's influential works (Ravenscroft 2009, 134–135).⁶ Regarding the second point— FP_2 renders the meanings of all mental terms holistically interdependent, but given that each and every user of such terms is unlikely to subscribe to exactly the same set of platitudes for each and every term, most of us turn out to speak rather different languages when seemingly using the very same mentalistic words (Fodor and Lepore 1992). In sum, FP_2 is definitely not an uncontroversial proposal about the semantics of mental terms. Turning now to FP_3 , the psychological hypothesis—few would commit themselves to a straightforward form of theory-theory these days. The information that is used in mindreading is not necessarily structured like a theory, nor is the cognitive mechanism involved necessarily one of detached theorizing.⁷ It is also questionable whether mental state ascription plays a central role in day-to-day social cognition at all (Ratcliffe 2007). Thus, if we focused merely on folk psychology

⁵ FP_2 and FP_3 map onto a helpful distinction between *external* and *internal* folk psychology made by Stich and Ravenscroft (1994, 63): “External accounts either collect or systematize the intuitively recognizable generalizations of commonsense psychology, while internal accounts focus on the cognitive mechanism that underlies our ability to have those intuitions, to predict behavior, etc. [...] Internal accounts use the label ‘folk psychology’ for the knowledge structures that actually underlie [sociocognitive] skills.” Also, see Stich and Nichols (2003) for another parallel distinction between *platitude approach* and *mindreading approach* to defining folk psychology.

⁶ Then again, critique from the Kripkean causal theory of reference only works if the theoretical terms in question are *natural kind terms*, for which the theory is meant to apply. See e.g., (McGinn 1978) for a denial of that.

⁷ Regarding the first point, see e.g., (Hutto 2008) for a proposal that folk-psychological knowledge is embedded in narratives. Regarding the second point, see e.g., (Goldman 1993) or (Stich and Nichols 2003) for prominent simulationist alternatives and (Gallagher 2008) for an interactionist one.

in that third sense—i.e., a common knowledge-structure about mental states that underlies the bulk of our natural sociocognitive skills—we might be led to conclude that there is no such thing as folk psychology (Morton 2007). Yet even from this it would not follow that there is no folk psychology *in the first sense*.

Furthermore, while the idea behind FP_2 stems from the observation that we use mental concepts to predict and explain behavior, this observation itself does not stand or fall together with FP_2 . In other words—it might be the case that a term-introducing theory (analogous with those found in sciences) is not the best way to account for the meanings of the conceptual repertoire in FP_1 . Nevertheless, departure from this analogy with scientific theories would not change the fact that we *sometimes* do use the concepts like ‘belief,’ ‘desire,’ ‘intention,’ etc., to explain and predict human behavior. Examples of this practice are commonplace enough. This is granted by even those who insist that there is much more to folk psychology beyond its explanatory and predictive function (see e.g., McGeer 2007, 139). While we might indeed use these concepts for other purposes too,⁸ the fact that we *also* use them to explain and predict puts FP_1 *prima facie* in competition with other explanatory projects, notably those of psychology and neuroscience—thus raising the question of how it ultimately relates to these projects.

To conclude this section, let me recap the three main points once more. First, it is a fact that we humans, at least the healthy adult ones, use mentalistic concepts (i.e., FP_1), and that at least sometimes we use these concepts to explain or predict behavior. Second, this fact does not depend on whether FP_3 is an accurate psychological account of social cognition in general (FP_4) or mental state ascription in particular (FP_1), or whether FP_2 is an accurate account of the semantics of mentalistic vocabulary (FP_1). This means that regardless of the recent anti-theory-theoretic and more generally anti-intellectualist trends in social cognition research, genuine unresolved issues concerning the explanatory and predictive potential of common mental vocabulary still remain. Third, FP_1 is legitimately called folk psychology—just as our commonsensical categories of living things are legitimately summarized under the label ‘folk biology,’ regardless of what the underlying mechanisms of categorization turned out to be.

⁸ Some recent proposals for alternative functions of folk psychology besides explanation/prediction include: making vicarious speech acts (Gauker 2003), enabling moral judgments (Knobe 2007), shaping uniform behavior in societies (Zawidzki 2013), managing relationships (Bohl 2015). Often, these are accompanied by a disparaging attitude towards the perceived “standard view” which takes explanation and prediction to be the *central* functions. I have no quibble with this—even an occasional use of folk psychology as an explanatory tool would be enough to motivate the problems discussed in this paper.

3. The claim for autonomy

In the following, then, I will focus on the potential autonomy of FP_1 (henceforth simply folk psychology). What would it mean for folk psychology in that sense to be autonomous? Recall that here we are first and foremost after *explanatory* autonomy—roughly the idea that explanations provided using mentalistic vocabulary are in some important way independent of those couched in non-mentalistic terms. This is to be contrasted with *ontological* autonomy—roughly the idea that such independence reflects some deeper divide in the world, not merely in our ways for making sense of the world. I want to set the ontological questions aside for now.⁹

To begin, it is useful to distinguish between two senses of explanation, in which I will be relying on José Luis Bermúdez (2005). On the one hand, we have explanations for particular dated events or states in terms of other states and events. Why did the door open? Because John pushed it. Why did John open the door? Because he wanted to get some fresh air and thought that opening the door will provide him with some. Bermúdez calls these *horizontal* explanations. On the other hand, we can also try to explain those aforementioned explanations themselves, so to speak. Why is it the case that doors open when you push them—what makes the latter fact explanatory of the former? Why is it that John will open the door when he wants to do so? Answers to such questions are properly called *vertical* explanations, for they involve descending from a certain *level* to another—certain explanatory regularities are explained in terms of other, lower-level regularities.¹⁰

What levels of explanation are there for human behavior? Given that we are physical beings, it is in principle possible to explain and predict our trajectories in purely physical terms. Given that we are biological systems, it is possible to do it in terms of biological mechanisms. Given that we are endowed with a sophisticated control system, i.e., networks of nerve cells, it can be done in terms of transmission of neural signals. Presuming that we are properly described as complex information processing devices, it can be done in terms of complex computational procedures. Now, regardless of what other intermediate levels there are, at the other end of this spectrum we find explanations that are couched in familiar terms of mental states that

⁹ I also aim to steer clear of the issues pertaining to reductionism as it is classically discussed in the philosophy of science. Reduction is most often understood as a relation between two *theories*, yet I do not want to commit to the view that mental concepts (FP_1) are necessarily embedded in a theory-like structure (FP_2) that could be considered for a wholesale reduction. Again, I find that mentalistic vocabulary can be seen as explanatory without assuming anything like a folk-psychological theory.

¹⁰ I will treat explanation and prediction as symmetrical, while acknowledging that this standard view has its dissenters (see e.g., Andrews 2003).

are attributed to rational agents. This is the level of folk psychology. Following an oft-cited distinction from Dennett (1969), we can call this latter level *personal* and all the lower levels *subpersonal*.¹¹

With this general picture in place, we can start assembling the claim for autonomy (see Fig. 1). Consider an event that is eligible for personal-level explanation, such as John intentionally opening the door. On the one hand, we could provide a horizontal personal-level explanation for it. We could cite certain mental states (M in Fig. 1) such as John's desire to get some fresh air and his belief that by opening the door there will be some fresh air, which would then explain why the action (A in Fig. 1) of John opening the door came about. On the other hand, we could describe and explain the event at various subpersonal levels. For instance, we might recount how certain patterns of John's neural activity (C1 in Fig. 1) led to a series of muscle movements (E1 in Fig. 1). Or even further down the levels, we could show how a series of chemical reactions in John's body (C2 in Fig. 1) led to certain other chemical reactions (E2 in Fig. 1). Such subpersonal explanations can—at least in many cases¹²—be connected by vertical explanations. This consists in explaining higher-level causal regularities in terms of lower-level causal regularities. For instance, it is possible to explain many of the regularities that hold on the aforementioned neural level with regularities that hold on the molecular level. This particular type of vertical explanation is well within the reach of contemporary science; others might be as of yet not, but are in principle achievable. Yet an autonomy theorist would maintain that the relation between the *personal level* and various subpersonal levels is altogether different. This is because those two styles of horizontal explanation—unlike any two levels of subpersonal explanation—are in principle incommensurable.

¹¹ There is some debate over the exegesis of Dennett's original distinction (Hornsby 2000; Drayson 2012; Drayson 2014). It is not easy to find a distinguishing feature for what is intuitively the personal level and not altogether clear what Dennett himself took it to be. However, I need not concern myself with exegetical issues here—thus, I take personal-level explanations to simply consist of familiar folk psychological states attributed to persons. Nor do I need to decide here what (if anything) distinguishes the *subpersonal* from merely *impersonal*, as e.g., Hornsby (1997, 161) would insist.

¹² There is the further issue of whether the kind of subpersonal psychology that invokes technical, yet nevertheless *intentional* notions (such as 'mental representation') could also be autonomous from the lower, non-intentional levels of explanation. The particular case for autonomy of folk psychology that I will be discussing here would not apply for subpersonal intentional psychology, so I will refrain from delving into the matter any further.

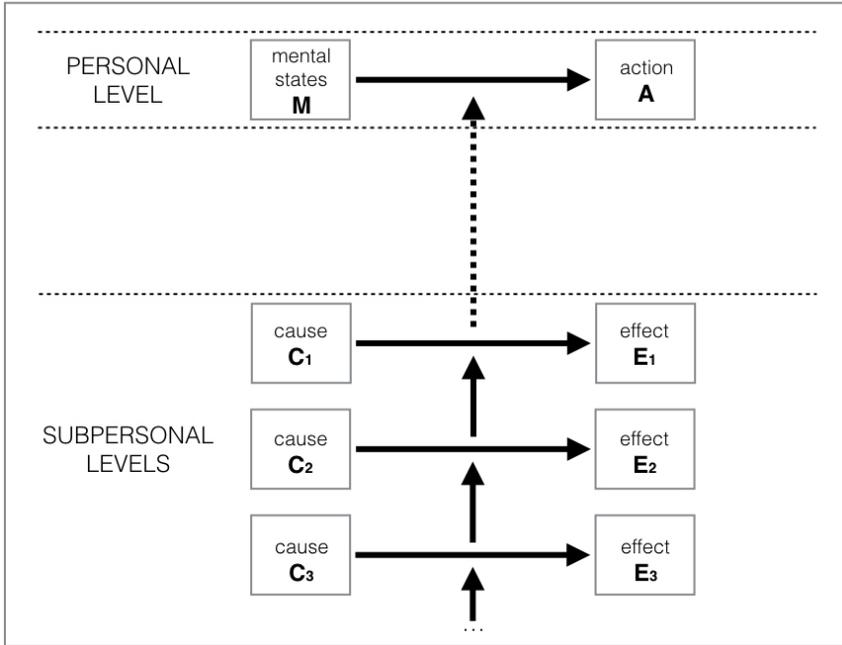


Figure 1. Three kinds of explanations for psychological phenomena

The claim for explanatory autonomy as I understand it, then, is the following: *there are no vertical explanations to be found for explaining the personal-level horizontal explanations* (cf. Bermúdez 2005, 31–51). In other words—the dotted vertical arrow in Fig. 1 above cannot be filled in for any explanation couched in folk-psychological terms. Or, in terms of the example mentioned above—pleas for vertical explanations such as “why is it that John will open the door when he wants to do so?” do not make any sense (cf. Dennett 1969, 93).

With this general idea of autonomy on the table, we can now ask why should it be taken seriously. Why is it that the personal-level explanatory statements themselves are not explainable? In the following, I will focus on only one, albeit classic argument—i.e., the one from normativity-*ex-rationality*. The version presented here draws on the views of Davidson (1974), McDowell (1998) and Hornsby (1997; 2000), although I will not be exploring the more nuanced details from any particular authors. Rather, the following should be read as a rational reconstruction of a certain key idea behind several slightly different philosophical positions.¹³

¹³ See (Bermúdez 2005, 45–51) for a nice summary of how the authors mentioned here—i.e., Davidson and McDowell—differ in their views pertaining to the autonomy of folk psychology.

While both the personal and subpersonal horizontal explanations on Figure 1 look superficially similar, they differ in one important respect—namely, the *way* in which the explanans explains the explanandum. At the personal level, a mental state (or a set of states) *rationalizes* an intentional action or possibly another mental state, while in subpersonal explanations, a state *causally* explains some other state.¹⁴ On closer inspection, then, these two kinds of horizontal explanation turn out to provide answers for rather different types of ‘why’-questions. In the personal case, we tell a story of why it *made sense* for an agent to think or do something—in light of certain principles of rationality we adhere to. In the subpersonal case, we tell a story of how a certain effect was produced by certain causes—in light of certain explanatory causal regularities.¹⁵ The causal regularities that we cite in explanations of the latter kind are *descriptive* statements about the world and clearly amenable to revision based on empirical inquiry. The principles of rationality, however, are *normative*—they do not originate from observed regularities, nor are they reducible to such in any straightforward way. Consider a typical example of such a principle: “if the subject desires that X and believes that Y-ing leads to X, then one should Y.” There is no obvious way for principles of that form to be sensitive to (i.e., revisable in light of) any empirical insights we might gain. As Davidson (1974, 231, my emphasis) puts it, “[in using personal-level explanation] we necessarily *impose* conditions of coherence, rationality and consistency. These conditions have no echo in physical theory.”

Thus, the argument consists of two claims: first, that all personal-level explanations invoke the normative principles of rationality; and second, that such normative principles cannot be understood in non-normative terms in any straightforward way (cf. Bermúdez 2005, 44). The second claim is essentially a recourse to the age-old Hume’s is-ought problem—let us grant for argument’s sake that the latter is on point. Regarding the first claim, David-

¹⁴ One might insist that a proper personal-level explanation also includes a causal component—the rationalizing states cited in a proper explanation need to be the causes of the action explained. Davidson’s (2001) metaphysical views, which consider mental states token-identical with physical states, can accommodate this while maintaining explanatory autonomy.

¹⁵ I am expressing this point in such a loose way (i.e., in terms of “certain explanatory causal regularities”) to maintain that the causal explanations on the subpersonal level are not necessarily of the deductive-nomological kind; i.e., such that the observed phenomenon is deduced from starting conditions and strict causal laws. It is generally agreed that the deductive-nomological account will not suffice in special sciences where there are no universal, exceptionless laws. Most popular contemporary account of explanation in neuroscience (and possibly psychology) is the mechanistic view (see e.g., Craver 2007). Thus, the placeholder “certain explanatory causal regularities” is meant to possibly include descriptions of underlying mechanisms.

son has extensively argued that without assuming the target to be rational, personal-level explanations cannot get off the ground. An example might help. Suppose that John is completely irrational—meaning that we could never count on him to behave rationally in light of any mental states we happened to ascribe to him. If there are no rational patterns to be expected, then there is no point in ascribing mental states, for such ascriptions would never explain *why* John happened to behave in one way or another. Thus, such ascriptions make sense only with the constitutive ideal of rationality in place. The flipside of this is that with every mentalistic ascription, we are making claims about how the subject *ought* to behave.

How does this affect the concepts of FP₁? On the autonomy theorists' view, these concepts have an exclusive purpose, given the considerations above—they “have their proper home in explanations of a special sort” (McDowell 1998, 389). Insofar they are exclusively part of this special self-sufficient explanatory framework, they are beyond scientific criticism, revision and elimination. This is the practical consequence of the autonomy of folk psychology.

Naturally, the conclusions drawn from the normativity of folk psychology and the alleged normative character itself have been much disputed. It is not possible to do justice to all the twists and turns of the burgeoning normativity debate here,¹⁶ so I will focus on one formidable, all-around critique of such normativity-to-autonomy moves, presented by Jonathan Knowles (2002). After all, my purpose in this section is merely to show that the case from normativity should not be dismissed out of hand. Also, dealing with a general objection will hopefully make the key idea behind the autonomy thesis clearer.

Knowles tackles three different elements that are often intertwined in the works of autonomy theorists—the *a priori* nature, normativity and holism of our mentalistic conceptual repertoire. Here I will focus on his case against normativity, although the general strategy that he follows is similar for all three aspects. Knowles (2002, 200) holds the autonomy theorists committed to the claim that folk psychology is in the business of giving *distinctly non-scientific* explanations.¹⁷ To establish this, they would have to demonstrate a distinctive feature which cannot be found in any scientific explanation. One possibility is to underline the apparent normativity of FP₁'s constitutive principles in the sense I have been discussing throughout this section. However,

¹⁶ See (McHugh and Whiting 2014) for one recent overview regarding the issues about the normativity of belief.

¹⁷ Although Knowles does not use the label ‘autonomy theorists’ for his targets, they are nevertheless the very same authors who subscribe to the autonomy of folk psychology in the sense I have been describing—so, I will adopt this convenient way of referring to them.

Knowles (2002, 212–214) argues that on closer inspection, the disanalogy between scientific explanation and commonsense rationalization is really not that strong. More specifically—on any coherent construal, the way in which constitutive principles figure in explanations of either kind is similar enough to undermine the disanalogy. He considers several alternative forms that a rational explanation could take (for ‘X believes B’):

- (1) X believes A
 X believes that if A, B
 If one believes A and if A, B, then one should (rationally)
believe B
 X believes B

[...]

- (2) X believes A
 X believes that if A, B
 X believes that if one believes A and if A, B, then one should
believe B
 X believes B

(Knowles 2002, 212)

How do norms figure in these explanations? In (1), the third premise is clearly a norm, yet it is unclear how it interacts with other premises to bring about the conclusion—thus, it either is redundant, or needs to be supplemented by a descriptive statement. In (2), there are no redundant premises, but then again, there are no norms directly involved in the inference; just a belief about a norm. If one did not presuppose a Davidsonian account of mental states (which at this point would mean begging the question), it would seem that (2) cites merely descriptive facts about X. Thus, it is structurally similar to subpersonal explanations in this sense. Furthermore, insists Knowles, this form of explanation could not count as a rationalizing explanation, for subjects might have some weird false beliefs about inferential rules, which could then lead them to irrational conclusions:

- (3) X believes A
 X believes that if A, B
 X believes that if one believes A and if A, B, then one should
believe C
 X believes C

(Knowles 2002, 213)

Thus, we are left with (4), which might strike as “an intuitively good explanation”:

- (4) X believes A
X believes that if A, B
 X believes B

(Knowles 2002, 213)

Here, some implicit form of normativity might be argued to originate from the fact that this explanation does not cite any causal laws. Yet there are also clearly non-normative explanations that do not explicitly cite any of such:

- (5) The Titanic crashed with an ice-berg
 The Titanic sank

(Knowles 2002, 213)

According to Knowles, there is no non-question-begging argument to be had for why (4) is principally different from the structurally similar (5). He concludes that folk-psychological form of explanation is not special due to some inherent normativity.

Although I am ultimately sympathetic with Knowles' aims, I do not think that the argument from normativity I have been discussing can be dismissed so easily. Firstly, it is not clear whether those defending autonomy would see personal-level explanations as aiming to explain statements with the logical form 'X believes B,' rather than 'X *should* believe B'—regardless of the surface grammar in any particular explanation. If the real explanatory target is the latter, then the premises in (1) would constitute a proper explanation:

- (1*) X believes A
 X believes that if A, B
If one believes A and if A, B, then one should (rationally) believe B
 X should believe B¹⁸

Secondly, I think (Knowles 2002, 213) is plainly wrong when he denies that in (3) the "explanandum is revealed to be as it rationally ought to be." When we learn of X's other beliefs, including the really weird one featured in the third premise, his belief in C is immediately cast in a rational light. Rationality does not entail truth. Thus, the fact that (2) is structurally analogous with (3) does not in itself count against (2) being a rationalizing explanation, for (3) could perfectly well be one.

Thirdly and perhaps most importantly, even if we grant that (1) and (1*) are in some sense defective and are thus forced to agree that personal-level explanations are structurally similar to subpersonal explanations, its consequences for the autonomy of folk psychology are unclear. Even granting that

¹⁸ This target could subsequently be transformed into a useful empirical explanation or prediction by adding the auxiliary premise that people usually behave in a rational manner.

the broader case Knowles brings against the disanalogy of folk-psychological and non-folk-psychological explanations is convincing, I find that there is something amiss with the way he sets up the whole problem. (Knowles 2002, 222) is dedicated to attacking the claim that “FP-explanations are of a different character from ordinary scientific explanations,” by which he actually seems to mean the claim that they are structurally distinct. But this is beside the point for an autonomy theorist. The difference that makes a difference, so to speak, is in the *content* of the constitutive principles rather than in the form of the explanation. Again: rational norms cannot be aligned with any set of causal regularities cited by psychologists and this renders vertical explanation of personal-level claims impossible. But this does not mean that rational norms have to figure in personal-level explanations in some unique formal way. Autonomy theory is not committed to the claim that folk psychology is unique in its autonomy—it is well compatible with the possibility that other levels of explanation (e.g., scientific disciplines such as biology) might turn out to be autonomous from lower levels too, if *their* constitutive principles turn out to be incompatible with the ones invoked on lower levels. So, folk psychology is not supposed to be distinct or special in *that* sense to be autonomous.

4. The limits of autonomy

I hope to have shown that the case from normativity to autonomy has at least some initial plausibility. This might explain why even philosophers of naturalistic persuasion are still regularly drawn to it, despite its somewhat anti-naturalistic and perhaps counterintuitive first appearances. Now, however, I want to show that this argument is not of much use to any champion of the personal level or folk psychology as such (e.g., Hornsby 2000). This is because at best, it applies only to a narrow segment of FP₁.

Recall what the argument from normativity entails: when we use FP₁ to explain or predict, then we are in the business of telling how things *should* be; i.e., ascribing mental states amounts to making normative claims. For example, if I ascribe to John the desire to get some fresh air and the belief that opening the door will provide him some, then by those very ascriptions I hold that, *ceteris paribus*, John *should* open the door. But is this *all* that mental ascriptions in explanatory contexts amount to? I propose a simple method for finding out. To be distinctly relevant for a normative claim, a mentalistic concept has to have some unique proprietary normative import—using the concept in an explanatory statement has to make a distinct contribution to what the statement claims that an agent should rationally do. Now, if there are distinctions between mentalistic concepts that do not differ in their normative contribution, then it is natural to suppose

that the distinction is in place for *descriptive* rather than normative considerations.¹⁹ The strategy, then, is to see whether there are any central distinctions in folk psychology that do not entail a difference in the normative dimension. Any such distinction is not subject to autonomy of the kind discussed above.

This requires a bit of elaboration. Let us introduce the term ‘*rational profile*’ for the set of everything that is/was rational for a subject to do (in the broadest possible sense) in light of her present mental states. Again—the particular flavor of autonomy theory under consideration here holds that folk psychological vocabulary is distinctly in the business of constituting such rational profiles. From this, a certain kind of autonomy follows for that vocabulary: what makes statements such as ‘S tried to bring about Q, because she wanted that P and believed that P leads to Q’ true cannot be explained in terms of lower-level regularities, for they themselves are explanatory only due to background assumptions about rationality and such principles “have no echo” on lower levels. A thorough-going autonomy theorist of folk psychology would claim that as a result of that, *all* personal-level concepts that figure in explanatory contexts have a distinct function of explicating (in conjunction with the principles of rationality) the rational profile of a subject, rather than simply *describing* her in a coarse-grained manner. In other words, on the extreme autonomy view, these concepts lack any content beyond the mark they leave on the rational profile. Thus, subpersonal-level talk cannot give us any insight into the phenomena that are invoked in this special task of explicating ones rational profile. Yet is this plausibly the case with all or even most of the mental state concepts? If this were so, then one would think that all or most of these concepts differ in what they contribute to the rational profile. If they do not and there are only a few normative distinctions among them, then what is the purpose of this abundance of mentalistic concepts? Barring the non-serious possibility that mankind has gradually introduced these numerous distinctions between otherwise synonymous concepts for simply aesthetic reasons such as, say, a desire for diversity, we could well suppose in that case that most of these concepts serve an additional function of *describing* the subject. But then a pair of, say, normatively synonymous propositional attitude concepts C₁ and C₂ in ‘S did X, because she C₁-d that Y and C₂-d that Z’ *does* plausibly have content beyond the mark they leave on the rational profile, and thus the statement indeed

¹⁹ While describing and evaluating might not be the only functions that a term could have (think of, e.g., various performatives), they are the two plausible ones for such linguistic items as the names of mental states. Thus it is plausible to assume that the lack of a normative function indicates a descriptive one.

has content which is in principle compatible with (and thus explainable by) subpersonal-level talk.

We can begin to chip away the alleged normativity of FP_1 by first considering purely phenomenal states. Interestingly, Dennett (1969, 93) first introduced the notion of personal level by using the example of pain. Now, an autonomy theorist is likely to hold that pain can be a perfectly good normative reason—that is, being in pain clearly affects what is rational for an agent to do.²⁰ However, it is also quite clear that being a normative reason cannot be the whole story about pain. Furthermore, it is easy to think of other phenomenal states which do not come packed with such well-recognizable motivational force. All sorts of vague, ambiguous and undefined feelings would be an example.

At this point, the autonomy theorist is likely to concede that weird phenomenal states are not what his theory is meant for. Yet weird phenomenal states occupy merely the fringes of folk psychology anyway—a case in point is that FP_1 does not contain many concepts for referring to particular kinds of vague, ambiguous feelings with no clear normative import. Instead, the autonomy is usually claimed for the true core of FP_1 : the concepts for various propositional attitudes, such as beliefs, desires, memories, perceptions, emotions, etc. *These* are the concepts that are employed in their proprietary normative explanations with their distinctive normative weight. Or so says the autonomy theorist.

But are they? Thinking of various distinctions among propositional attitude concepts, it is not clear at all whether they trace the lines of unique normative contributions. Naturally, there is a distinction with key normative import between cognitive (i.e., belief-like) and conative (i.e., desire-like) states. Desiring that *P* affects the rational profile clearly in a very different way than believing that *P*. Yet zooming in any further in either family of mental states, the uniqueness of contributions from the perspective of rationality becomes much more dubious. Compare, for example, the following cognitive ascriptions:

(B) *S* believes that *P*

(M) *S* remembers that *P*

Believing and remembering are undoubtedly among the central concepts of folk psychology—but is there a normative distinction here? Do these two ascriptions lead to different results with regard to what is rational for *S*? It would seem that not—what matters for the rational profile is simply that *S*

²⁰ This is unproblematic if one subscribes to some perceptual or motivational theory of pain.

takes it to be the case that P, of which both (B) and (M) are an instance. To repeat the example from earlier—it makes no difference for assessing whether John's act of opening the door was rational, whether he believes or remembers that opening the door provides him with some fresh air.

Maybe this particular choice of example is unfortunate, however, as it invites the following reply: there definitely is a difference in the rational profiles of (B) and (M) subjects, because remembering (unlike belief) is commonly held to be factive—remembering that P entails truth of P (Bernecker 2010, 36–39). Thus, (B) and (M) clearly differ in their normative contribution—namely, given that (M) presupposes the truth of P, it warrants a lot more than (B). Fair enough, let us suppose that this is the case—but then consider a triad that is even more telling:

(M) S remembers that P

(P) S perceives that P

(K) S knows that P

(M), (P) and (K) all ascribe cognitive states that are commonly held to be factive (Williamson 2002). Does it make any difference for the rational profile in the situation, whether the subject is said to remember, know or perceive that P? On the common understanding of these mentalistic terms, the answer seems to be no.²¹ *Ceteris paribus*, the subject is rationally warranted to think and act in an identical way in all three cases. According to a common conception, what (M) and (P) bring to table are, roughly, concepts for two distinctive *ways* of knowing—and given that those ways are not distinctive in the normative sense, these concepts have to differ in their descriptive content. This should not be much of a shock and might even sound rather trivial. But noting it explicitly serves to underline the fact that the argument

²¹ Someone might dispute this assessment with the following example: Kim and Tim want to go to a certain shop. Both can be said to know its location—Kim lives nearby and sees the shop just down the street from her window; Tim has been to the shop once before and remembers the address. Even though both have factive propositional attitudes with the content 'The shop is located at X,' it might seem that Kim's knowledge is somehow more secure than Tim's and consequently, different paths of action would be rational for Kim and Tim (e.g., for Tim, it would be rational to double-check the address before going, but not so for Kim).

I think that the proponent of such an objection faces a dilemma. Either he uses the verbs 'remembers' and 'sees' in the standard sense, according to which both are factive states produced by a reliable faculty, or he does not. If he uses the verbs in the standard sense, then the intuition that Kim and Tim have an unequal amount of justification and different rational profiles is unwarranted. If he does not use the verbs in the standard sense, then his objection does not concern the distinctions made in FP₁. Thanks to Francesco Orsi for pointing out this objection.

for autonomy presented in section 3 does not apply with at least some of the concepts in these particular distinctions—and here we are dealing with some concepts that lie at the very heart of FP_1 .

Someone might object at this point: why would a lack of normative distinctions eliminate the normativity—and *ipso facto* the autonomy—of any concepts? Could not a set of *concepts* be normative even when there are no normative *distinctions* among them?²² Indeed—(M), (P) and other propositional attitude ascriptions *do* have some normative import in the context of a rationalizing explanation, after all (never mind that it is often of exactly the same kind). Yet the claim is not that there is *no* normative component to these propositional attitude concepts nor that the above strategy could reveal any of these concepts to be wholly non-normative. Rather, what is at issue is whether making a mark on the rational profile is *all* that these concepts are about, for *that* would be the guarantee of their autonomy given the argument in the section 3. But the above considerations provide strong evidence that this just could not be all that many central mentalistic concepts are about. In addition to their influence on the rational profile, they clearly have specific descriptive content, which is perfectly eligible to figure in causal explanations of the kind that are compatible with subpersonal-level talk.

Take for instance the statement: ‘S did A, because S perceived that P’ It includes the attribution of a state that belongs firmly in FP_1 —i.e., a perception. On the one hand, attributing this perception to S means altering her rational profile. If the statement is an instance of a successful personal-level explanation, then it has revealed to us a part of why it made sense for S to do A. Why this is so— i.e., why this rationalization holds—is plausibly beyond any further vertical explanations which cite merely psychological regularities in S’s cognitive machinery. Insofar this is all that the concept of ‘perception’ is supposed to do, it remains autonomous. On the other hand, however, attributing a perception to S also means *describing* S as being in a certain state—and not just in some or other state which happens to be token-identical with profile-altering rational force of perception, but in a *specific* state as defined by the descriptive part of ‘perception.’ For instance, it is a state that differs from merely having knowledge, perhaps in that it involves exercising perceptual modalities; has certain temporal and content constraints, etc. In other words, the descriptive part of ‘perception’ picks out a state that can figure in standard causal explanations. Thus, the argument from normativity does not suffice for ruling out that there is *something* to be said about the connection of S’s perception and consequent behavior on the subpersonal level.

²² Thanks to an anonymous reviewer for bringing up this point.

We could continue this exercise with various other propositional attitude names beyond belief, desire, perception and memory, but I think I have said enough to illustrate the main point: there is a wide range of importantly distinct folk psychological concepts, yet merely a few normative distinctions among them, meaning that a large part of folk psychological machinery cannot be in place for the special purpose of rationalizing. My diagnosis is that the autonomy-from-normativity claim I have been discussing might be better suited for a certain regimented core of a particular folk psychological practice—i.e., for something along the lines of Dennett’s ‘intentional systems theory,’ which includes just the stripped-down versions of belief and desire (Dennett 1987, 58).²³ FP₁, on the other hand, contains much more than simply the abstract incarnations of belief, desire and consequent intentional action. Insofar the bulk of its content does not constitute a system of normative distinctions, FP₁ as a whole is not covered by the argument from normativity in any meaningful sense.

5. Conclusion

Let me recap the topics of this paper. From about mid-1980s to mid-1990s, there was a huge debate about the status of folk-psychological conceptual inventory as an explanatory tool (see e.g., the papers collected in Greenwood 1991). At some point, this discussion rapidly dwindled down—not necessarily because a consensus was achieved, but more likely because it, as so often happens in philosophy, was overshadowed by new and more pressing concerns. To be more specific—as the idea of mindreading subserved by a psychologically real theory-representation became increasingly unpopular, philosophers became more and more involved in discussing the other plausible mechanisms of social cognition and less interested in the vaguely theory-like edifice manifested in our everyday mentalistic conceptual framework. As I have tried to show in this paper by distinguishing the four independent meanings of ‘folk psychology,’ the question about the explanatory potency of ordinary mental state concepts has not simply dissolved because of any developments with regard to FP₂, FP₃ or FP₄.

There is a certain segment of philosophers who keenly insist (and probably many others who tacitly agree) that folk-psychological explanations are autonomous from lower-level explanations. A commonly invoked reason for this is that mental state ascriptions are tied to a constitutive ideal of ratio-

²³ “[...] The first new theory, intentional system theory, is envisaged as a close kin of, and overlapping with, such already existing disciplines as decision theory and game theory, which are similarly abstract, normative, and couched in intentional language. It borrows the ordinary terms ‘belief’ and ‘desire’ but gives them a technical meaning within the theory” (Dennett 1987, 58).

nality (Davidson 1974). On this view, ‘John believes that P’ and ‘Sally intends to Q’ are in effect normative rather than descriptive statements—they tell us something about how John and Sally should behave *qua* rational agents. I have argued that if a statement conveys purely normative content in this sense, then it is indeed plausibly considered as autonomous—but the former is not the case with most of the explanations given in folk psychological terms. Most folk psychological concepts clearly have content beyond their effects on the rational profile and thus, the normativity argument provides no reason for why they could not figure in typical causal explanations.

To end with a conjuncture—I suspect that the partly negative result reached with the normativity argument generalizes for other strategies that try to determine the status of the whole folk psychology or personal level in one brush. As it is widely recognized by now, folk psychology—even in the restricted sense of FP₁—is a hugely variegated phenomenon and probably cannot be neither wiped out nor vindicated with a single decisive maneuver.

Acknowledgments

The first draft of this paper was presented at a departmental colloquium in University of Tartu—I would like to thank those present for their helpful discussion and comments. The work on this paper was supported by the Estonian Science Foundation grant ETF9117, the Centre of Excellence in Estonian Studies (European Union, European Regional Development Fund) and is related to research project IUT20-5 (Estonian Ministry of Education and Research).

Bibliography

- Andrews, K. (2003). Knowing mental states: The asymmetry of psychological prediction and explanation, in Q. Smith and A. Jokic (eds), *Consciousness: New Philosophical Perspectives*, Oxford University Press, Oxford, pp. 201–219.
- Bermúdez, J. L. (2005). *Philosophy of Psychology: A Contemporary Introduction*, Routledge, New York.
- Bernecker, S. (2010). *Memory: A Philosophical Study*, Oxford University Press, Oxford.
- Bohl, V. (2015). We read minds to shape relationships, *Philosophical Psychology* 28: 674–694.

- Craver, C. F. (2007). *Explaining the Brain: Mechanisms and the Mosaic Unity of Neuroscience*, Oxford University Press, Oxford.
- Davidson, D. (1974). Psychology as philosophy, in S. C. Brown (ed.), *Philosophy of Psychology*, Macmillan Press, London, pp. 41–52.
- Davidson, D. (2001). Mental events, *Essays on Actions and Events*, 2nd edn, Clarendon Press, Oxford, pp. 207–225.
- Dennett, D. C. (1969). *Content and Consciousness*, Routledge, London.
- Dennett, D. C. (1987). *The Intentional Stance*, The MIT Press, Cambridge, MA.
- Drayson, Z. (2012). The uses and abuses of the personal/subpersonal distinction, *Philosophical Perspectives* **26**: 1–18.
- Drayson, Z. (2014). The personal/subpersonal distinction, *Philosophy Compass* **9**: 338–346.
- Fodor, J. A. and Lepore, E. (1992). *Holism: A Shopper's Guide*, Wiley-Blackwell, Cambridge, MA.
- Gallagher, S. (2008). Inference or interaction: Social cognition without precursors, *Philosophical Explorations* **11**: 163–174.
- Gauker, C. (2003). Attitudes without psychology, *Facta Philosophica* **5**: 239–256.
- Godfrey-Smith, P. (2005). Folk-psychology as a model, *Philosophers' Imprint* **5**: 1–16.
- Goldman, A. I. (1993). The psychology of folk psychology, *Behavioral and Brain Sciences* **16**: 15–28.
- Goldman, A. I. (2000). Folk psychology and mental concepts, *Protosociology* **14**: 4–25.
- Gopnik, A. and Meltzoff, A. N. (1997). *Words, Thoughts, and Theories*, A Bradford Book, Cambridge, MA.
- Greenwood, J. D. (ed.) (1991). *The Future of Folk Psychology: Intentionality and Cognitive Science*, Cambridge University Press, Cambridge.
- Hornsby, J. (1997). *Simple Mindedness: In Defence of Naive Naturalism in the Philosophy of Mind*, Harvard University Press, Cambridge, MA.
- Hornsby, J. (2000). Personal and sub-personal: A defence of Dennett's early distinction, *Philosophical Explorations* **3**: 6–24.
- Hutto, D. D. (2008). *Folk Psychological Narratives: The Sociocultural Basis of Understanding Reasons*, The MIT Press, Cambridge, MA.

- Jackson, F. and Pettit, P. (1990). In defence of folk psychology, *Philosophical Studies* 59: 31–54.
- Knobe, J. (2007). Folk psychology: Science and morals, in D. D. Hutto and M. Ratcliffe (eds), *Folk Psychology Re-assessed*, Springer, Dordrecht, pp. 157–174.
- Knowles, J. (2002). Is folk psychology different?, *Erkenntnis* 57: 199–230.
- Lewis, D. K. (1972). Psychophysical and theoretical identifications, *Australasian Journal of Philosophy* 50: 249–258.
- McDowell, J. H. (1998). *Mind, Value, and Reality*, Harvard University Press, Cambridge, MA.
- McGeer, V. (2007). The regulative dimension of folk psychology, in D. D. Hutto and M. Ratcliffe (eds), *Folk Psychology Re-assessed*, Springer, Dordrecht, pp. 137–156.
- McGinn, C. (1978). Mental states, natural kinds and psychophysical laws I, *Proceedings of the Aristotelian Society, Supplementary Volumes* 52: 195–236.
- McHugh, C. and Whiting, D. (2014). The normativity of belief, *Analysis* 74: 698–713.
- Medin, D. L. and Atran, S. (1999). *Folkbiology*, MIT Press, Cambridge, MA.
- Mölder, B. (2011). Normatiivuse mitu tahku, *Studia Philosophica Estonica* 4: 42–82.
- Morton, A. (2007). Folk psychology does not exist, in D. D. Hutto and M. Ratcliffe (eds), *Folk Psychology Re-assessed*, Springer, Dordrecht, pp. 211–222.
- Ratcliffe, M. (2007). *Rethinking Commonsense Psychology: A Critique of Folk Psychology, Theory of Mind and Simulation*, Palgrave Macmillan, Hampshire and New York.
- Ratcliffe, M. and Hutto, D. D. (2007). Introduction, in D. D. Hutto and M. Ratcliffe (eds), *Folk Psychology Re-assessed*, Springer, Dordrecht, pp. 1–22.
- Ravenscroft, I. (2009). Is folk psychology a theory?, in J. Symons and P. Calvo (eds), *The Routledge Companion to Philosophy of Psychology*, Routledge, chapter New York, pp. 131–147.
- Sellars, W. S. (1956). Empiricism and the philosophy of mind, in H. Feigl and M. Scriven (eds), *Minnesota Studies in the Philosophy of Science*, University of Minnesota Press, chapter Minneapolis, pp. 243–329.

- Stich, S. P. and Nichols, S. (2003). Folk psychology, in S. P. Stich and T. A. Warfield (eds), *The Blackwell Guide to Philosophy of Mind*, Blackwell Publishing, Malden, MA, pp. 235–255.
- Stich, S. P. and Ravenscroft, I. (1994). What is folk psychology?, *Cognition* **50**: 447–468.
- Stueber, K. R. (2009). The ethical dimension of folk psychology?, *Inquiry* **52**: 532–547.
- Williamson, T. (2002). *Knowledge and Its Limits*, Oxford University Press, Oxford.
- Zawidzki, T. W. (2013). *Mindshaping: A New Framework for Understanding Human Social Cognition*, The MIT Press, Cambridge, MA.